



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 161820

TO: Celine Qian  
Location: REM-2A64&2C70  
Art Unit: 1636  
Thursday, August 11, 2005  
  
Case Serial Number: 09/913878

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512  
  
edward.hart@uspto.gov

### Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: 09/913878 Examiner #: 78910 Date: 8/8/95  
 Art Unit: 1636 Phone Number: 2-0777 Serial Number: 09/913878  
 Mail Box and Bldg/Room Location: 2464 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of invention: Isolation and characterization of an Crassa Silencing gene & uses thereof

Inventors (please provide full names):

Giuseppe et al.

Earliest Priority Filing Date: 2/22/1999

CRFE

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

8045 NA

Please search SEQ ID NO: 1 from 2447-6652  
 (commercial & interference database)

## STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>8/9/95</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>8/11/95</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. Review Time _____	Fulltext _____	Sequence Systems <u>04</u>
Clerical Prep. Time _____	Patent Family _____	WWW/Internet _____
Online Time _____	Other _____	Other (specify) _____

**This Page Blank (uspto)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 13:42:12 ; Search time 17518 Seconds

(without alignments)  
11633.899 Million cell updates/sec

Title: US-09-913-878A-1\_COPY\_2447\_6652

Perfect score: 4206

Sequence: 1 atgaaccttacttactctag.....tcacagggaatgcgattat 4206

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_bcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sls:\*

12: gb\_gy:\*

13: gb\_un:\*

14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4206	100.0	4206	8 NCR133528	AJ133528 Neurospor
2	4206	100.0	8045	6 AX034425	AX034425 Sequence
3	4202.8	99.9	4206	6 CQ803434	CQ803434 Sequence
4	3078	73.2	3078	6 CQ803436	CQ803436 Sequence
5	55.2	1.3	125020	9 AF429315	AF429315 Homo sapi
6	53.8	1.3	6688	8 AF468822	AF468822 Diaporthe
7	52.4	1.2	2000	6 AX655393	AX655393 Sequence
8	51.8	1.2	2000	6 AX655393	AX655393 Sequence
9	51.6	1.2	1164	6 A10243	A10243 S. murinus D
10	51.4	1.2	3524	8 AF443073	AF443073 Phomopsis
11	50.8	1.2	125020	9 AF429315	AF429315 Homo sapi
12	50	1.2	1610	1 AB038265	AB038265 Streptomy
13	50	1.2	1952	1 AF170068	AF170068 Streptomy
14	49	1.2	3731	6 ARI45904	ARI45904 Sequence
15	49	1.2	3731	8 LERDRP	Y10403 L. esculentu
16	48.6	1.2	11492	1 AB012149	AB012149 Xanthomon
17	48.4	1.2	1551	1 SSPY1518	Y15518 Streptomyce
18	48.4	1.2	1750	1 AF257177	AF257177 Streptomy
19	47	1.1	19601	1 AB070944	AB070944 Streptomy

20	47	1.1	299425	1 AP005049	AP005049 Streptomy
21	46.8	1.1	3674	8 AY574374	AY574374 Nicotiana
22	45.8	1.1	3800	1 SAR416377	SA416377 Streptomy
23	45.8	1.1	6204	8 AF411016	AF411016 Neurospor
24	45.6	1.1	2710	6 AX059555	AX059555 Sequence
25	45.4	1.1	6987	8 AY049072	AY049072 Diaporthe
26	44.6	1.1	194083	10 AC132253	AC132253 Mus muscu
27	44.6	1.1	1349742	1 BX572090	BX572090 Pichloro
28	44.2	1.1	162959	2 AC019255	AC019255 Homo sapi
29	44.2	1.1	178217	9 AC022080	AC022080 Homo sapi
30	44.2	1.1	197866	2 AC068293	AC068293 Homo sapi
31	44	1.0	720	6 CQ718786	CQ718786 Sequence
32	44	1.0	1318	6 AF025354	AF025354 Homo sapi
33	44	1.0	1831	6 AX405860	AX405860 Sequence
34	44	1.0	1919	8 BC017798	BC017798 Homo sapi
35	44	1.0	1959	6 CQ492849	CQ492849 Sequence
36	44	1.0	2149	9 AK057696	AK057696 Homo sapi
37	44	1.0	75440	9 AP000786	AP000786 Homo sapi
38	44	1.0	118218	2 AP000670	AP000670 Homo sapi
39	44	1.0	190739	2 AP001361	AP001361 Homo sapi
40	44	1.0	192229	9 AP000943	AP000943 Homo sapi
41	43.8	1.0	34445	9 AC006291	AC006291 Homo sapi
42	43.8	1.0	75214	8 BX842597	BX842597 Neurospor
43	43.8	1.0	77292	8 BX842596	BX842596 Neurospor
44	43.8	1.0	114918	9 AC005188	AC005188 Homo sapi
45	43.8	1.0	153667	9 AC109439	AC109439 Homo sapi

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
NCR133528	NCR133528	Neurospora crassa gde-1 gene, partial.	AJ133528	AJ133528.1	GI:4803726	gde-1 gene; RNA-dependent RNA polymerase.	Neurospora crassa	Neurospora crassa	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.	1	Cogoni, C. and Macino, G.	Gene silencing in neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase	Unpublished	2 (bases 1 to 4206)	Cogoni, C.
								Direct Submission							
								Submitted (08-MAR-1999)							
								Ematologia, Universita di Roma La Sapienza, Viale Regina Elena							
								324, 00161 Rome, ITALY							
								Location/Qualifiers							
								1. .4206							
								/organism="Neurospora crassa"							
								/mol_type="genomic DNA"							
								/db_xref="taxon:5141"							
								1. .4206							
								/gene="gde-1"							
								1. .>4206							
								/gene="gde-1"							
								/note="putative"							
								/function="gene silencing"							
								/codon_start=1							
								/product="RNA-dependent RNA polymerase"							
								/protein_id="CA842634.1"							
								/db_xref="GI:4803727"							
								/db_xref="GOA:Q9Y7G6"							
								/db_xref="UniProt/TREMBL:Q9Y7G6"							
								/translation="WNPIPRKNSPVEIINRIANDYNGLCQVADTTLTPRRKSL							
								AESEDDGRHDKIYRALNPLVWRDSDLSNOAEANFFLEAKAASSNWPKHADDTLP							
								WSKEPPRAATAGQOMALQTVLLEVTIANFMFPNNPTPRTIGRTLSGSGLSRPTSTWT							

KRDEPANTVFADPPKSLTRSATGPPIHGAALPLKPPDVNTGSKPSLESENANC  
TKRAGLSDNVAAAAPVPIASALDKVPRRANTRDPTAGHRADVDSDTSQ  
GTSYSSVFSACRNOOSITQSPFAPSOPEKRPVDAITFEAGHLESPKRTTS  
HINDQPLSSSQGTSFSTYSESPSGGAIPEPSRNGLASSESASQVAPV  
VVARLNIWPKPKMLHEAPLAVAMEVTLPMCKYDLDESLGLYDMSWARTV  
TDIMKTLXRLDAPRGKPEPKPNDVDFVMTGPNESKSAVUSLADYNPNSPVA  
PLIVYKLPLMFEGKCLTRFGRFPELLIPSTSVSVPPVYSKQGABEVIO  
MLTWGSHLVGRQKRAFPAKDAGRKRLREFQLAEDPKI1KERVHFPAETG1TFEP  
DVEKTRSVPAEPEVQRTERKVSQMLDMLQJDNNTQPHKLFR1QJLSTYAI  
MTLEPHQIRHKITDLSPSGEYMDGVMSRVAKR1RDVLGLDLSASQVRG  
SAKMWIVIDDVDTGDEDM1ETYP5QRKMECDPVDKHORTLEVSVAELASQINLO  
LPLVEDBARDKVKKROAIGRLIND1ROPSQOKHALNREVPWVSESRATV  
SHGRVPLAIGPDSQEBTLNPLMNSGDPKKQKTLQDIANDLQKRDITLKSXNITV  
GRSAV1YMIADPFWGLNEENEHVGFSSKEFDEESFTLLSDCVLVARSPHFPDIO  
RRAVFPRELHSLKDV11FSTGDPVPLAKLSGGDYGDMAWCMDEI1VDGFVNAEM  
PLBPLSLRYLKDKTTFKOLMAHSHTGSAKECTTYMIQSFHPLQAPNPLGCTNY  
KERLCYINNSVSNKPAL1SSLVGNLVDQSKOG1VEPNEA5OALRELLGALSTLP  
MYKSDMLGGEPTHI1DYLKFSIARPAIDKELEA1FNANKAKADSDP1PVRVNOY  
SYTFPKEISDKSRSSALLFTYTLKNR1GEVBEKE1GR1VKNKEMRSDP1PVRVNOY  
EKWCAT1PEAMDKSGANYDSKVLRLLESLADEMENTMAL1RSLASTF1KLYYHSPK  
VWQMAQRQLAV1KAQMTSRPEGA1PALMTAFMYAGLMPDKKFTQYVARLEGGSEYP  
DPEYEVLEGGDDPFG1GFTNGDY"

## ORIGIN

Query Match	100.0%	Score 4206	DB 8	Length 4206
Best Local Similarity	100.0%	Pred. No. 0		
Matches 4206	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy 1	ATGAACCTTATTA	CTCTAGAGAAAGAGAAATGACCCCGTCGAGAAATCATTAACCGGCTC	60	
Db 1	ATGAACCTTATTA	CTCTAGAGAAAGAAATGACCCCGTCGAGAAATCATTAACCGGCTC	60	
Qy 61	AATAAGACTACAA	CTTGAGGCTTCAGTGTGTGCGACAGACAACTTCAACCCCGCACGCG	120	
Db 61	AATAAGACTACAA	CTTGAGGCTTCAGTGTGTGCGACAGACAACTTCAACCCCGCACGCG	120	
Qy 121	CGGAAGAGCTGCG	CGGAGAGTGAAGAGATTTCGCTGCGCATGACAGATTTACAGAGCC	180	
Db 121	CGGAAGAGCTGCG	CGGAGAGTGAAGAGATTTCGCTGCGCATGACAGATTTACAGAGCC	180	
Qy 181	CTGAACCTTTCT	CTACTGCGGAGAGATGACTCCCTGAACAGGAGAGCAACCTTCTTC	240	
Db 181	CTGAACCTTTCT	CTACTGCGGAGAGATGACTCCCTGAACAGGAGAGCAACCTTCTTC	240	
Qy 241	ATCGAGGCCAAAG	CTGCGAGCTCGAACTGGGTTGCCAAAGCCACGCGCACCTTGACAGC	300	
Db 241	ATCGAGGCCAAAG	CTGCGAGCTCGAACTGGGTTGCCAAAGCCACGCGCACCTTGACAGC	300	
Qy 301	CTTCCGCTGCTCA	AGAACTCCCGCGCGCTACTCTGCGCGCCCAACATGGGCTTTCAG	360	
Db 301	CTTCCGCTGCTCA	AGAACTCCCGCGCGCTACTCTGCGCGCCCAACATGGGCTTTCAG	360	
Qy 361	ACTGTGTGCTCG	AGTGTGCTTAATAGGTTATGCGACCTCCCAATAACACACAGGTGCA	420	
Db 361	ACTGTGTGCTCG	AGTGTGCTTAATAGGTTATGCGACCTCCCAATAACACACAGGTGCA	420	
Qy 421	ACGTTTGGCAAA	CTTAAGCGGCCCAAGTGGCTTGAACCGGCCCAACCTTACCAACACC	480	
Db 421	ACGTTTGGCAAA	CTTAAGCGGCCCAAGTGGCTTGAACCGGCCCAACCTTACCAACACC	480	
Qy 481	AAACGCAAGATGA	CGCCGCAATGTCACTTTCGCTGATCGCGCCAAACGCTGTTACT	540	
Db 481	AAACGCAAGATGA	CGCCGCAATGTCACTTTCGCTGATCGCGCCAAACGCTGTTACT	540	
Qy 541	CGCTCTGCAAG	GTCTCTATTCAGCGGCGCGCGGATACCCCTTAAGTTCCCGGATCCA	600	
Db 541	CGCTCTGCAAG	GTCTCTATTCAGCGGCGCGCGGATACCCCTTAAGTTCCCGGATCCA	600	
Qy 601	GTCGAATACCG	GTCTCAACACATCTCTCGAGATGAGATCTCAATCAGTGCACCAAG	660	
Db 601	GTCGAATACCG	GTCTCAACACATCTCTCGAGATGAGATCTCAATCAGTGCACCAAG	660	
Qy 661	CGGCGCAAGGCA	AGTGTCTGATTAATGTTGCGCTGCGCGCGCGCGCGCGCGCTTATT	720	

Db 661	CGGCGCAAGGCA	AGTGTCTGATTAATGTTGCGCTGCGCGCGCGCGCGCGCGCTTATT	720
Qy 721	GGCAGGCGCTT	GGACAAGGATACCGACTCGAAGGATGCGCAATGAGAGATCCACGCGC	780
Db 721	GGCAGGCGCTT	GGACAAGGATACCGACTCGAAGGATGCGCAATGAGAGATCCACGCGC	780
Qy 781	ACAAGTCATAG	ACAGCGGACAGCGGACAGGTGATTCCTTTGATACATCTCAAGGCACTTCTAT	840
Db 781	ACAAGTCATAG	ACAGCGGACAGCGGACAGGTGATTCCTTTGATACATCTCAAGGCACTTCTAT	840
Qy 841	GGTTCCAGTGT	CTTTCAGCGCTTCCCGTTCACATACAGAGCATACCCAGATAGTTTGGAG	900
Db 841	GGTTCCAGTGT	CTTTCAGCGCTTCCCGTTCACATACAGAGCATACCCAGATAGTTTGGAG	900
Qy 901	GCTCTCTCTT	CAAGCCCAAGAGAAAGGCGCGCTGTGATGCGACGCTTTGAGGCTGGA	960
Db 901	GCTCTCTCTT	CAAGCCCAAGAGAAAGGCGCGCTGTGATGCGACGCTTTGAGGCTGGA	960
Qy 961	CACCTGATTA	GATCTCTAGCAAAAGAAAGAAACAACCAAGTCCCAATGATTAACAAGCCC	1020
Db 961	CACCTGATTA	GATCTCTAGCAAAAGAAAGAAACAACCAAGTCCCAATGATTAACAAGCCC	1020
Qy 1021	CTTTCATCTCT	CTTCCAGGCTGAAACTTGTGAGACCTTACTATGATGCTTCCAACT	1080
Db 1021	CTTTCATCTCT	CTTCCAGGCTGAAACTTGTGAGACCTTACTATGATGCTTCCAACT	1080
Qy 1081	TCCGCGCGCA	GGGCGCAATTTCCGAGCGGAGTGTGCTCAAAATGAGCTGGCTCGAGCGAA	1140
Db 1081	TCCGCGCGCA	GGGCGCAATTTCCGAGCGGAGTGTGCTCAAAATGAGCTGGCTCGAGCGAA	1140
Qy 1141	GAAGGCGCTGA	TCAGGTTCAATGCCCGGCTGTGACGCTCGGCTGAGAAAT	1200
Db 1141	GAAGGCGCTGA	TCAGGTTCAATGCCCGGCTGTGACGCTCGGCTGAGAAAT	1200
Qy 1201	ATTTGGCCGAA	ATTTCCCAATGCTCAACGAAAGCTCTCTGCTGTGTGATGGAGATT	1260
Db 1201	ATTTGGCCGAA	ATTTCCCAATGCTCAACGAAAGCTCTCTGCTGTGTGATGGAGATT	1260
Qy 1261	ACGAGACTTTA	TGACATGCAAGTGAACCTTGGAGAGCGAGAGCTGGGCTTAAGTAC	1320
Db 1261	ACGAGACTTTA	TGACATGCAAGTGAACCTTGGAGAGCGAGAGCTGGGCTTAAGTAC	1320
Qy 1321	GACCTTCTGT	GCTACCGCGCGAGTGTCAAGATATCTGGAAGACTCTTACCGGCTT	1380
Db 1321	GACCTTCTGT	GCTACCGCGCGAGTGTCAAGATATCTGGAAGACTCTTACCGGCTT	1380
Qy 1381	GATGCTTTCG	TGATAACCTTTCCAGAAAGCGCGCCCAACGATGTTCCGACCGCA	1440
Db 1381	GATGCTTTCG	TGATAACCTTTCCAGAAAGCGCGCCCAACGATGTTCCGACCGCA	1440
Qy 1441	ATGACGGGCA	ACTTTGAGAGCAAGTGAAGTGCCTGTGTTCTCTGCTGTTCTAGACTAC	1500
Db 1441	ATGACGGGCA	ACTTTGAGAGCAAGTGAAGTGCCTGTGTTCTCTGCTGTTCTAGACTAC	1500
Qy 1501	AATCCGACA	CACTGCTACTGCGCCCTTTACTTGTGAAGCTGAAGCCGCTCATGTTTC	1560
Db 1501	AATCCGACA	CACTGCTACTGCGCCCTTTACTTGTGAAGCTGAAGCCGCTCATGTTTC	1560
Qy 1561	GAGCAGGCGT	GATCAACCCGTCGGTTCGGTTCCTGATAGGTTTTTTCAGAGATCTTTATA	1620
Db 1561	GAGCAGGCGT	GATCAACCCGTCGGTTCGGTTCCTGATAGGTTTTTTCAGAGATCTTTATA	1620
Qy 1621	CCGTGCTTGA	GCACACAGCCCAAGTGTACCGCGGTGTGACAAACAAGGTGCG	1680
Db 1621	CCGTGCTTGA	GCACACAGCCCAAGTGTACCGCGGTGTGACAAACAAGGTGCG	1680
Qy 1681	GTCGAAGATCA	TTCAGTGTGCTCACGATGGGCAACATTTCTTGTAGCGCGCAATGG	1740
Db 1681	GTCGAAGATCA	TTCAGTGTGCTCACGATGGGCAACATTTCTTGTAGCGCGCAATGG	1740
Qy 1741	CGGCTTTCTT	CGCCCAAGTGCAGATACAGAAACCTTCTCAGGAGTTCCAGCTCCGC	1800
Db 1741	CGGCTTTCTT	CGCCCAAGTGCAGATACAGAAACCTTCTCAGGAGTTCCAGCTCCGC	1800



Dh 1741 CGCGCTTTCTTCCGCAAGATGCCGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 1800  
Qy 1801 GCGGAGACCCGAAACCATCATCAAGAGAGAGTCCACTTCTTGGCGAGACCGGCATT 1860  
Db 1801 GCGGAGACCCGAAACCATCATCAAGAGAGAGTCCACTTCTTGGCGAGACCGGCATT 1860  
Qy 1861 ACGTTCGACCTGATGTGTTCAAGACGAGATCTGTCTCCGCGAGAGAACTGTAGAG 1920  
Db 1861 ACGTTCGACCTGATGTGTTCAAGACGAGATCTGTCTCCGCGAGAGAACTGTAGAG 1920  
Qy 1921 CAACGACCGAGTTCAAAAGTTAGTCAAAATGCTGACGCTGCTCTGCAACTGTACACAC 1980  
Db 1921 CAACGACCGAGTTCAAAAGTTAGTCAAAATGCTGACGCTGCTCTGCAACTGTACACAC 1980  
Qy 1981 ACTTGGAGCGGACCTCAAGTTGTCTCCGTAATCCAGCTCGGTCTGAGTAAGACATAT 2040  
Db 1981 ACTTGGAGCGGACCTCAAGTTGTCTCCGTAATCCAGCTCGGTCTGAGTAAGACATAT 2040  
Qy 2041 GCCATTATGACATTGGAGCCTCACCAAGTACAGACACCAAGACCGATCTTCTTCACT 2100  
Db 2041 GCCATTATGACATTGGAGCCTCACCAAGTACAGACACCAAGACCGATCTTCTTCACT 2100  
Qy 2101 TCAGGCACTGGCGAAGTATGATGATGACGCTGAGCGGCATGTCCGAAAGCTGGCCAG 2160  
Db 2101 TCAGGCACTGGCGAAGTATGATGATGACGCTGAGCGGCATGTCCGAAAGCTGGCCAG 2160  
Qy 2161 AGGATAGCGCATGTCTCGTGTGGGTGATGCTGCGCTCTGCTGCAAGGGCGGTTGGT 2220  
Db 2161 AGGATAGCGCATGTCTCGTGTGGGTGATGCTGCGCTCTGCTGCAAGGGCGGTTGGT 2220  
Qy 2221 TCGGCGCAAGGAAATGTGGGTTATCGACGTTGACGACACAGCGATGAGATTTGGATCGAG 2280  
Db 2221 TCGGCGCAAGGAAATGTGGGTTATCGACGTTGACGACACAGCGATGAGATTTGGATCGAG 2280  
Qy 2281 ACATACCCTGTCACGCGCAAGTGGAAATGCGACTTCTGTGTAATAACATCAACGTACCTC 2340  
Db 2281 ACATACCCTGTCACGCGCAAGTGGAAATGCGACTTCTGTGTAATAACATCAACGTACCTC 2340  
Qy 2341 GAAGTCGAGGCGGTGCTTCTGAACTGAAGTCACTGCTCTCAACCTACAGCTGTACCT 2400  
Db 2341 GAAGTCGAGGCGGTGCTTCTGAACTGAAGTCACTGCTCTCAACCTACAGCTGTACCT 2400  
Qy 2401 GTCTCGAAGATAGAGCCAGGACGAAAGTGAAGATGCGCCAGGCAATCGGTGACCTT 2460  
Db 2401 GTCTCGAAGATAGAGCCAGGACGAAAGTGAAGATGCGCCAGGCAATCGGTGACCTT 2460  
Qy 2461 ATCAACGATTTGCAACGACAGTTCAGCGAGCAAAAGCATGCTTGAATCGCCAGTGGAA 2520  
Db 2461 ATCAACGATTTGCAACGACAGTTCAGCGAGCAAAAGCATGCTTGAATCGCCAGTGGAA 2520  
Qy 2521 TTTTGGCAATGGGTTTTCAGAGTTATTCAGTGGCGCACTCGAGTCAAGCAGGCGCT 2580  
Db 2521 TTTTGGCAATGGGTTTTCAGAGTTATTCAGTGGCGCACTCGAGTCAAGCAGGCGCT 2580  
Qy 2581 GTGCTTTTCTTGTGGGCTTCTGACAGTCAAGAGAGACACTGAACTTCTTATGTAAC 2640  
Db 2581 GTGCTTTTCTTGTGGGCTTCTGACAGTCAAGAGAGACACTGAACTTCTTATGTAAC 2640  
Qy 2641 AGTGGGTTGATCCCAAGCAAAAGTACTTGCAGAATCGCTGGGATCTTCAAAAG 2700  
Db 2641 AGTGGGTTGATCCCAAGCAAAAGTACTTGCAGAATCGCTGGGATCTTCAAAAG 2700  
Qy 2701 CGGAAATGTGACAGTTGAAGTCCAGCTGAACATCCGTGCGGTCAATGCAATTAAT 2760  
Db 2701 CGGAAATGTGACAGTTGAAGTCCAGCTGAACATCCGTGCGGTCAATGCAATTAAT 2760  
Qy 2761 TACATGATGTGCGAATTTCTGGGGTGTGCTTGAAGAAATGAGTTCAATGTGCAATTC 2820  
Db 2761 TACATGATGTGCGAATTTCTGGGGTGTGCTTGAAGAAATGAGTTCAATGTGCAATTC 2820  
Qy 2821 TCAAAAGTTCAAGACGAGAGAGAGTCTTTTACACTCTATCGACTGTGATGTCTCGTG 2880  
Db 2821 TCAAAAGTTCAAGACGAGAGAGAGTCTTTTACACTCTATCGACTGTGATGTCTCGTG 2880

Qy 2881 GCGGATCCCCAGCCCATTTCCCTAGATATCCAAAGGCTGAGCAGTCTTCAAGCA 2940  
Db 2881 GCGGATCCCCAGCCCATTTCCCTAGATATCCAAAGGCTGAGCAGTCTTCAAGCA 2940  
Qy 2941 GAGCTCCACAGTCTCAAGATGATATCATCTTCTTACTTAAGAGATGTACCGCTTGGT 3000  
Db 2941 GAGCTCCACAGTCTCAAGATGATATCATCTTCTTACTTAAGAGATGTACCGCTTGGT 3000  
Qy 3001 AAGAAGTATCTGTGAGAGCTACGAGGAGATATGGCCGTGCTGTGGATCCGGAG 3060  
Db 3001 AAGAAGTATCTGTGAGAGCTACGAGGAGATATGGCCGTGCTGTGGATCCGGAG 3060  
Qy 3061 ATGTCGATGGTTTCTGCAATGCGGAAATGCTCTGAGGCCGACCTGTCTAGTACTTA 3120  
Db 3061 ATGTCGATGGTTTCTGCAATGCGGAAATGCTCTGAGGCCGACCTGTCTAGTACTTA 3120  
Qy 3121 AAGAAGCAAAACGACTTTCAAACTTAATGAGCTTCAACGAGCGGCTCAGCGGCC 3180  
Db 3121 AAGAAGCAAAACGACTTTCAAACTTAATGAGCTTCAACGAGCGGCTCAGCGGCC 3180  
Qy 3181 AAGAAGCAAGTACATACGATATGATTCAGAAAGCTTCAATTTGCGCTCGACGCCAAC 3240  
Db 3181 AAGAAGCAAGTACATACGATATGATTCAGAAAGCTTCAATTTGCGCTCGACGCCAAC 3240  
Qy 3241 TTTCTGGGCAATGCACTAACCTAACAAAGAGGCTGTGATCAATCAATAGTGTCT 3300  
Db 3241 TTTCTGGGCAATGCACTAACCTAACAAAGAGGCTGTGATCAATCAATAGTGTCT 3300  
Qy 3301 AACAGCGGCGCATCATTTCTTAAGTTCACTGTGGGAAACCTCGTGCATCAGACCAACAA 3360  
Db 3301 AACAGCGGCGCATCATTTCTTAAGTTCACTGTGGGAAACCTCGTGCATCAGACCAACAA 3360  
Qy 3361 GGTATGTCTTTTAAAGAGCAAGCTGGGCTCAATTTGGGTAGGGAATGCTTTGGCGGTGCA 3420  
Db 3361 GGTATGTCTTTTAAAGAGCAAGCTGGGCTCAATTTGGGTAGGGAATGCTTTGGCGGTGCA 3420  
Qy 3421 TTTGCTCCCTTCTGACCCCAATGTACAGAGCGAAGTGGTCTGGGGCGGAGAGCTTACC 3480  
Db 3421 TTTGCTCCCTTCTGACCCCAATGTACAGAGCGAAGTGGTCTGGGGCGGAGAGCTTACC 3480  
Qy 3481 CACATTATGATACCTGAATTCCTCATGCGCAGGCTCGATGTGCAAGAACTGGAA 3540  
Db 3481 CACATTATGATACCTGAATTCCTCATGCGCAGGCTCGATGTGCAAGAACTGGAA 3540  
Qy 3541 GCTTTCACAAATGCTATGAAAGCGGCGCAAGATACAGAAAGCGGCTCATTTTGGGAT 3600  
Db 3541 GCTTTCACAAATGCTATGAAAGCGGCGCAAGATACAGAAAGCGGCTCATTTTGGGAT 3600  
Qy 3601 CCGGATCTGCTTCTTACTACAGTTCAGGTTCTTCAAGAGATTGAGCAAGTGGCTTCC 3660  
Db 3601 CCGGATCTGCTTCTTACTACAGTTCAGGTTCTTCAAGAGATTGAGCAAGTGGCTTCC 3660  
Qy 3661 GCACTGCTATTCAAGACTCTGAAGAAACCGTATCGGCGAAGTCAAGAAATATGCGAGG 3720  
Db 3661 GCACTGCTATTCAAGACTCTGAAGAAACCGTATCGGCGAAGTCAAGAAATATGCGAGG 3720  
Qy 3721 TTTGCTCAAAACCAAGAGATGAGAGACAGCAAGAACCCCTTCTGCGCTCAACAG 3780  
Db 3721 TTTGCTCAAAACCAAGAGATGAGAGACAGCAAGAACCCCTTCTGCGCTCAACAG 3780  
Qy 3781 GTTTATGAAAAATGGGCGGCATCAGCGCTGAAGCGGATGCAAAATCGGAGCAAAATTAT 3840  
Db 3781 GTTTATGAAAAATGGGCGGCATCAGCGCTGAAGCGGATGCAAAATCGGAGCAAAATTAT 3840  
Qy 3841 GATTTCAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGAGACCGTGAATGAATACA 3900  
Db 3841 GATTTCAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGAGACCGTGAATGAATACA 3900  
Qy 3901 TGGGCAATGTGAGGGCTTACAGGCTTTCAACTGTACTACCAAGAGACCCCAAGTTC 3960  
Db 3901 TGGGCAATGTGAGGGCTTACAGGCTTTCAACTGTACTACCAAGAGACCCCAAGTTC 3960

QY 3961 GTGTGGCAGATGGGCGGCAGACAGCTCGCTGATCATTTAAGCGCAGATGACGACGACCCC 4020  
DB 3961 GTGTGGCAGATGGGCGGCAGACAGCTCGCTGATCATTTAAGCGCAGATGACGACGACCCC 4020  
QY 4021 GGTGAAGCGCCCGCGGGTGTGATGACCGCGTTGATGATGCGGGCTTGATGCGCGGATTAAG 4080  
DB 4021 GGTGAAGCGCCCGCGGGTGTGATGACCGCGTTGATGATGCGGGCTTGATGCGCGGATTAAG 4080  
QY 4081 AAGTTTACGAAGCAGTATGTGGCCAGGCTGGAAGGCGCATGATCGAGTACCTGATCCG 4140  
DB 4081 AAGTTTACGAAGCAGTATGTGGCCAGGCTGGAAGGCGCATGATCGAGTACCTGATCCG 4140  
QY 4141 GAGGTCTATGAAGTGTGGCGGATGATGATTTGATGATGATTTGTTTCAAGGGAATGCG 4200  
DB 4141 GAGGTCTATGAAGTGTGGCGGATGATGATTTGATGATGATTTGTTTCAAGGGAATGCG 4200  
QY 4201 GATTAT 4206  
DB 4201 GATTAT 4206  
RESULT 2  
AX034425 8045 bp DNA linear PAT 22-SEP-2000  
LOCUS Sequence 1 from Patent WO0050581.  
DEFINITION AX034425  
ACCESSION AX034425 GI:10303130  
VERSION AX034425.1 GI:10303130  
KEYWORDS  
SOURCE Neurospora crassa  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
REFERENCE  
1 Cogoni, C. and Macino, G.  
Isolation and characterization of a n. Crassa silencing gene and  
uses thereof  
Patent: WO 0050581-A 1 31-AUG-2000;  
UNIV ROMA (IT) ; COGONI CARLO (IT) ; MACINO GIUSEPPE (IT)  
Location/Qualifiers  
1. 8045  
/organism="Neurospora crassa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5141"  
2447..6655  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC10121.1"  
/translation="MNPITPRKRNSPVEEIIINLNDYNLGIQVADTTLTPHREKL  
AESDEDEGRDIXYRALNPLVYRKDSLNQAEANPTEAKAASNNWPKAHADPTLP  
WSKEPPAAATAGQOMALQVLLEVLNKFMPNNPTGRTTGRSLSGSPSTNT  
KRDEPANVTFADEPKRSLSRATGPIHGAALPLKRPDEVNTGSKRPSLEENLNC  
TKRAKGLSDNVAAAAPVPVIAALDKVTRRHANATDVAEAFADQVDSFDTSQ  
GTYSGVSFASCHNOSTTOSFEAPSPORERKPYDAVTEAGHLIESPSKRTTSS  
HINOPLSSSSGESTSTSTYESFPSSGEGAIPEPSASGLARSESASOVYNAAP  
VVAARLNTWPKPKMLHEAPLANEYVTLPMHCKYDLEDESLGKTPDSMTAADY  
TDLWKLRLDAFRGKPFEPKPNDFVLTMTNPFSSKSAVLSVLDPNPSPLA  
PLVTLKLPLEFGQCRLTRFGPDRFELLIPSTPSVSPVSKDGAEBVIO  
MLTNGHSLVGRQMRAPFAKADGKRLREFQLEADPKFIIEKRVHAFETDITRP  
DVFETRSVVAEPEORTTEFKVOMLDMLOJDNNTWOHLTLFRIOGLSKYTAI  
MTLEPHOIRHKTDLISPSGTGVMDGVBMRSVAKRIADVLGQVPSAOGSRG  
SAGKMYIVDPTDDEDMITTPYPSQKMECDPYDKORTLEYSVASELSKAGLNLQ  
LVLEDEKARKVKRQAIIGRLINDLRQCSBOKHALNRFVFRWVYESYSRARI  
SHGRVPLLAGLPDSQEBTLNPLMNSGDFPKQYLODIADLQKRCDTLKSUNIRV  
GSAVAYIMADFWGVLEENHVHGFSGKPEDEESFTLLSDCVLVARSAFHPDIO  
RLAVAFKPELSLADKVIIFSTKGDVPLAKLGGADYGDMAWCMPEIYDGVNANM  
PLEPDI.SRYLKKDKTTEPKOLMASHGSGSAKEQOTTYDMIOKSPHPLQAPFLGCMNY  
KEELCYNNVSVKPAIILISLIYANVDSKOSIVPRKASMAIDREBLGALSLEDP  
MYSDSMWLGSEPTTHIIDYLFKSIARADLEKLEAHNAKAKADKEDGAFHVPDIA  
SYTFPEKEISDKSRSSALLFTLKNRIGVEKEKGRLVKNKEMRDKDPVAVNVY  
EKWCAITPEAMDGSGANYDSKVLRLLESLADREMTWALNLAFAFLYHKSXPK  
VMWAGROLAYIKQMTSRPGEAPALMTAFMYAGLMPDKKFTKQYVARLEGDSSEYP  
DEPVEYVLGDDDFDGI.GFTNGDY"

ORIGIN  
Query Match 100.0%; Score 4206; DB 6; Length 8045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAACCCCTATTACTCTTAGAAGAGAAATAGCCCCCTGAGGAATCTTAAACGGCTC 60  
DB 2447 ATGAACCCCTATTACTCTTAGAAGAGAAATAGCCCCCTGAGGAATCTTAAACGGCTC 2506  
QY 61 AATTAACGACTACCAACTGGGCGCTCCAGTGTGACAGACAACTGACCCCGCACCG 120  
DB 2507 AATTAACGACTACCAACTGGGCGCTCCAGTGTGACAGACAACTGACCCCGCACCG 2566  
QY 121 CGAAGAGCTGGCCGAGAGTGAAGAGATTTTCGTGCGCATGACAAAGATCTTACAGAGCC 180  
DB 2567 CGAAGAGCTGGCCGAGAGTGAAGAGATTTTCGTGCGCATGACAAAGATCTTACAGAGCC 2626  
QY 181 CTGAACCTTTCTCTACTGCGGAAAGATGACTCCCTTAACCGACGACAACTTCTTC 240  
DB 2627 CTGAACCTTTCTCTACTGCGGAAAGATGACTCCCTTAACCGACGACAACTTCTTC 2686  
QY 241 ATGAGAGCCAAAGCTGCGAGCTCGAACTGAGTGGCCAAAGCCGACGCGATGACAG 300  
DB 2687 ATGAGAGCCAAAGCTGCGAGCTCGAACTGAGTGGCCAAAGCCGACGCGATGACAG 2746  
QY 301 CTTCCGTGATCCAAAGAACCTCCCGCGCGCTTACTGCGGCGCAACATGGGATTTGAG 360  
DB 2747 CTTCCGTGATCCAAAGAACCTCCCGCGCGCTTACTGCGGCGCAACATGGGATTTGAG 2806  
QY 361 ACTGCTTCTGAGAGTGTCTTAATAGTTTATGCGCACTTCCCAATPAACACAGGTGCA 420  
DB 2807 ACTGCTTCTGAGAGTGTCTTAATAGTTTATGCGCACTTCCCAATPAACACAGGTGCA 2866  
QY 421 ACCTTTGGCAGAACTTAAAGCGGCCCAAGTGGCTGAGCGGCCCAACTCTTACCAACACC 480  
DB 2867 ACCTTTGGCAGAACTTAAAGCGGCCCAAGTGGCTGAGCGGCCCAACTCTTACCAACACC 2926  
QY 481 AAACGCAAGATGAGCGCCGCAATGTCACTTTGCTGATTCGCGCCAAACGCTTGTGACT 540  
DB 2927 AAACGCAAGATGAGCGCCGCAATGTCACTTTGCTGATTCGCGCCAAACGCTTGTGACT 2986  
QY 541 CGCTTCGCCACAGGTCTCTTATTCACGCGCGGGCGATACCCCTTAAGTTCCCGCATCCA 600  
DB 2987 CGCTTCGCCACAGGTCTCTTATTCACGCGCGGGCGATACCCCTTAAGTTCCCGCATCCA 3046  
QY 601 GTGAATACCGGTTCCAAAGCAACATCTCTGAGAGTGAATCTCAATGAGTGAACAAG 660  
DB 3047 GTGAATACCGGTTCCAAAGCAACATCTCTGAGAGTGAATCTCAATGAGTGAACAAG 3106  
QY 661 CGGCGCAAGGGCAAGCTGTGTGATTAATGTTGCGCGTGCAGCGCCGCGCTGCTATT 720  
DB 3107 CGGCGCAAGGGCAAGCTGTGTGATTAATGTTGCGCGTGCAGCGCCGCGCTGCTATT 3166  
QY 721 GCGAGGCGTTTGGACAAAGTACCGACTCGAAGGCAATGCCAATPACAGAGATCCCAAGGGG 780  
DB 3167 GCGAGGCGTTTGGACAAAGTACCGACTCGAAGGCAATGCCAATPACAGAGATCCCAAGGGG 3226  
QY 781 ACAGGTCAATGACGAGCGGACGAGTGAATCTCTTTGATATCTTCAAGGCACTTCTTAT 840  
DB 3227 ACAGGTCAATGACGAGCGGACGAGTGAATCTCTTTGATATCTTCAAGGCACTTCTTAT 3286  
QY 841 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGCACTTACCAAGTATGTTGAG 900  
DB 3287 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGCACTTACCAAGTATGTTGAG 3346  
QY 901 GCTCCCTCTTCAAGCGCCAGAGAAAGCGGCGCTGAGTGAAGCCAGCTTTAGAGCTGGA 960  
DB 3347 GCTCCCTCTTCAAGCGCCAGAGAAAGCGGCGCTGAGTGAAGCCAGCTTTAGAGCTGGA 3406  
QY 961 CACTTGAATGAGTCTCTTACGAAAGAAACAACAATGCTCCACATAGATTAACGAGCCC 1020  
DB 3407 CACTTGAATGAGTCTCTTACGAAAGAAACAACAATGCTCCACATAGATTAACGAGCCC 3466

Qy 1021 CTTTCATGCTGCTCCAGGGGTAACCTTGCTTACGACCTTACTATGAGTGTTCCAACT 1080  
Db 3467 CTTTCATGCTGCTCCAGGGGTAACCTTGCTTACGACCTTACTATGAGTGTTCCAACT 3526  
Qy 1081 TCCGCGCGCGAGGGCGCAATCCCGAGCCGATCGCTCAATGAGACTGGCTCGAGCGAA 1140  
Db 3527 TCCGCGCGCGAGGGCGCAATCCCGAGCCGATCGCTCAATGAGACTGGCTCGAGCGAA 3586  
Qy 1141 GAAAGCGCTGATCTCAGATTCAAGTTCAATGCCCGGTGGTTGCACTCGGCTGAATAAT 1200  
Db 3587 GAAAGCGCTGATCTCAGATTCAAGTTCAATGCCCGGTGGTTGCACTCGGCTGAATAAT 3646  
Qy 1201 ATTTGGCGGAAATTTCCCAATGGCTACAGAACTCCTCGCTGTTGCAATGGGAATT 1260  
Db 3647 ATTTGGCGGAAATTTCCCAATGGCTACAGAACTCCTCGCTGTTGCAATGGGAATT 3706  
Qy 1261 ACCAGACTTTTATGCACTGCAAGTAGACTTGGAGAAGAGAGCCCTGGCTAAAGTAC 1320  
Db 3707 ACCAGACTTTTATGCACTGCAAGTAGACTTGGAGAAGAGAGCCCTGGCTAAAGTAC 3766  
Qy 1321 GACCTTCTGCTCTACCGCGCGGATGTCAAGATATCTGAAAGACTCTTACCGGCTT 1380  
Db 3767 GACCTTCTGCTCTACCGCGCGGATGTCAAGATATCTGAAAGACTCTTACCGGCTT 3826  
Qy 1381 GATGCTTTCGCTGTAACCTTTCCAGAAAGCGGCCAAGAGCTGTTGTAAGCGCA 1440  
Db 3827 GATGCTTTCGCTGTAACCTTTCCAGAAAGCGGCCAAGAGCTGTTGTAAGCGCA 3886  
Qy 1441 ATGACGGGCACTTTGAGAGCAAGTAGTGCCTGTTCTCTGCTGTTCTAGACTAC 1500  
Db 3887 ATGACGGGCACTTTGAGAGCAAGTAGTGCCTGTTCTCTGCTGTTCTAGACTAC 3946  
Qy 1501 AATCCGGAACACTGCGCTACTGCGCCCTTTACTTGTGAAGCTGAGCCGCTCATGTTT 1560  
Db 3947 AATCCGGAACACTGCGCTACTGCGCCCTTTACTTGTGAAGCTGAGCCGCTCATGTTT 4006  
Qy 1561 GAGAGGGCTGTGCACTCAACCGTGGTGGTGGTCTGTAAGTTTTCGAGATCCTTATA 1620  
Db 4007 GAGAGGGCTGTGCACTCAACCGTGGTGGTGGTCTGTAAGTTTTCGAGATCCTTATA 4066  
Qy 1621 CCGTGCCTTACGAGCACCAGGCCCAAGTGTACCGCGGTGTACAGCAAAACAACAGATGCG 1680  
Db 4067 CCGTGCCTTACGAGCACCAGGCCCAAGTGTACCGCGGTGTACAGCAAAACAACAGATGCG 4126  
Qy 1681 GTCGAAGAATCATCACTGAGTGTCAAGATGGGCAACTTCTTTGTAAGCCGCAATGG 1740  
Db 4127 GTCGAAGAATCATCACTGAGTGTCAAGATGGGCAACTTCTTTGTAAGCCGCAATGG 4186  
Qy 1741 CGCGCTTCTTCCGCAAAAGATGCGGGAATACAGAAACCTTCAGGGAATTCAGCTCCG 1800  
Db 4187 CGCGCTTCTTCCGCAAAAGATGCGGGAATACAGAAACCTTCAGGGAATTCAGCTCCG 4246  
Qy 1801 GCCGAGACCCGAAACCATCATCAAGAGAGAGTCACTTCTTTGCGGAGCCGCAATT 1860  
Db 4247 GCCGAGACCCGAAACCATCATCAAGAGAGAGTCACTTCTTTGCGGAGCCGCAATT 4306  
Qy 1861 AGCTTCCGACTGATGTGTTCAAGAGATCTGCTGTTCCGCGAGAGAACTGTAGAG 1920  
Db 4307 AGCTTCCGACTGATGTGTTCAAGAGATCTGCTGTTCCGCGAGAGAACTGTAGAG 4366  
Qy 1921 CAAGGAGCCGAGTTCAAGTTAGTCAATGCTGAGCTGGCTCTGCAATCGACAACAC 1980  
Db 4367 CAAGGAGCCGAGTTCAAGTTAGTCAATGCTGAGCTGGCTCTGCAATCGACAACAC 4426  
Qy 1981 ACTTGGACCGGCACTCTCAAGTTGTTCTCCGTAATCAGCTCGGTCTGAGTAAGCATAT 2040  
Db 4427 ACTTGGACCGGCACTCTCAAGTTGTTCTCCGTAATCAGCTCGGTCTGAGTAAGCATAT 4486  
Qy 2041 GCCATTATGACATTGAGGCTTCAAGATCAGACACCAAGACCGATCTTTCACTT 2100  
Db 4487 GCCATTATGACATTGAGGCTTCAAGATCAGACACCAAGACCGATCTTTCACTT 4546

Qy 2101 TCAGGCACTGGCGAAGTATGATGAAGTGTAGAGCCGCACTGTGCGAAGCGTGCCAAAG 2160  
Db 4547 TCAGGCACTGGCGAAGTATGATGAAGTGTAGAGCCGCACTGTGCGAAGCGTGCCAAAG 4606  
Qy 2161 AGATACGCGATGTTCTCGGTTTTGGTGTATGTGCCCTCTGCTGTGCAAGGCGGTTTTGT 2220  
Db 4607 AGATACGCGATGTTCTCGGTTTTGGTGTATGTGCCCTCTGCTGTGCAAGGCGGTTTTGT 4666  
Qy 2221 TCCGCCAAGGGAATGAGGGTTATCGACGTTGAGACACAGGCCATGAGATTGGATGGAG 2280  
Db 4667 TCCGCCAAGGGAATGAGGGTTATCGACGTTGAGACACAGGCCATGAGATTGGATGGAG 4726  
Qy 2281 ACATACCCGTCACAGCGCAAGTGGAAATGCGACTTCGTTGATAAATCAACGTAACCTC 2340  
Db 4727 ACATACCCGTCACAGCGCAAGTGGAAATGCGACTTCGTTGATAAATCAACGTAACCTC 4786  
Qy 2341 GAAGTCGAGCGTGCTTCTGAACTGAAATCAGCTGTGTCAACCTTACAGCTGTTACT 2400  
Db 4787 GAAGTCGAGCGTGCTTCTGAACTGAAATCAGCTGTGTCAACCTTACAGCTGTTACT 4846  
Qy 2401 GTCTGGAAGATAGACGACAGGCAAGAGTGAAGATGGCGCAAGCAATCGGTAACCGTCT 2460  
Db 4847 GTCTGGAAGATAGACGACAGGCAAGAGTGAAGATGGCGCAAGCAATCGGTAACCGTCT 4906  
Qy 2461 ATCAACGATTTGCAACGACAGTTCAGCGCAACAAAGCATCTTTGAATTCGCCAGTGA 2520  
Db 4907 ATCAACGATTTGCAACGACAGTTCAGCGCAACAAAGCATCTTTGAATTCGCCAGTGA 4966  
Qy 2521 TTTGCGCAATGGGTTTACGAGATTATTCAGTGGCGCAACTGAGTCAAGCACGCGCT 2580  
Db 4967 TTTGCGCAATGGGTTTACGAGATTATTCAGTGGCGCAACTGAGTCAAGCACGCGCT 5026  
Qy 2581 GTGCTTTTCTGCTGGGCTTACCTGACAGTCAAGAGAGACACGAACTTTGATGAGAC 2640  
Db 5027 GTGCTTTTCTGCTGGGCTTACCTGACAGTCAAGAGAGACACGAACTTTGATGAGAC 5086  
Qy 2641 AGTGGTTTCATCCCAAGAACAAAGTACTTGCACACATTCGCTGGATCTTCAAAAG 2700  
Db 5087 AGTGGTTTCATCCCAAGAACAAAGTACTTGCACACATTCGCTGGATCTTCAAAAG 5146  
Qy 2701 CGGAAATGTGACAGTTGAAATCCAGCTGAACATCCGTGTGATGATGATACAT 2760  
Db 5147 CGGAAATGTGACAGTTGAAATCCAGCTGAACATCCGTGTGATGATGATACAT 5206  
Qy 2761 TACATGATTCGCAATTTCTGCGGTGCTGAGGAAATGAGGTTCAATGCTGCTCC 2820  
Db 5207 TACATGATTCGCAATTTCTGCGGTGCTGAGGAAATGAGGTTCAATGCTGCTCC 5266  
Qy 2821 TCAAGTTTACAGACGAGAGAGTCTTTTACACTCTTATCGACTGTGATGTTCTCGTG 2880  
Db 5267 TCAAGTTTACAGACGAGAGAGTCTTTTACACTCTTATCGACTGTGATGTTCTCGTG 5326  
Qy 2881 GCGGATTCGCCAGCCATTTTCCCTAGTATATCCACGSGTTGAGCATCTTCAAGCCA 2940  
Db 5327 GCGGATTCGCCAGCCATTTTCCCTAGTATATCCACGSGTTGAGCATCTTCAAGCCA 5386  
Qy 2941 GAGCTTCAAGTCTCAAGATGTATATCTTCTCTCTAATAAGGAGATGTAACGCTTGCT 3000  
Db 5387 GAGCTTCAAGTCTCAAGATGTATATCTTCTCTCTAATAAGGAGATGTAACGCTTGCT 5446  
Qy 3001 AAGAAGCTATCTGGTGAAGACTACGAGCGGATATGCGCTGTGCTGGATCCGAG 3060  
Db 5447 AAGAAGCTATCTGGTGAAGACTACGAGCGGATATGCGCTGTGCTGGATCCGAG 5506  
Qy 3061 ATCTGATGTTTGTGTAATGCGAAATGCTCTGAGACCCGACTGTCTAGGTACTTA 3120  
Db 5507 ATCTGATGTTTGTGTAATGCGAAATGCTCTGAGACCCGACTGTCTAGGTACTTA 5566  
Qy 3121 AAGAAGCAAAAGCACTTTCAACAACTTATGGCTTCAACGCGCAAGGCTCAGCGGCC 3180  
Db 5567 AAGAAGCAAAAGCACTTTCAACAACTTATGGCTTCAACGCGCAAGGCTCAGCGGCC 5626  
Qy 3181 AAGAAGCAACTACATACGATATGATCAGAAGAGCTTCAATTTGCTTGCCTGACGCCAAC 3240

```
Db      ||| 5627 AAGAGAGAGACTATACATATATATCAGAAAGCTTCAATTCGCTCGACGCCAAC 5686
Qy      ||| 3241 TTCTTGGGAGATGTCACATCACTAACAAGAAAGCTGTGTTACATCAACAATAGTGTCT 3300
Db      ||| 5687 TTCTTGGGAGATGTCACATCACTAACAAGAAAGCTGTGTTACATCAACAATAGTGTCT 5746
Qy      ||| 3301 AACAAAGCCGGGCATCATTTCTTTAGTTCACTGGTGGAAACCTCGTCAATCAGAGCAACAA 3360
Db      ||| 5747 AACAAAGCCGGGCATCATTTCTTTAGTTCACTGGTGGAAACCTCGTCAATCAGAGCAACAA 5806
Qy      ||| 3361 GGTATTTCTTTTAAAGAGCAAGCTGGGCTCAATGGCTAGGAACTGGCTGGGCTGCA 3420
Db      ||| 5807 GGTATTTCTTTTAAAGAGCAAGCTGGGCTCAATGGCTAGGAACTGGCTGGGCTGCA 5866
Qy      ||| 3421 TTGCTCCCTCTGACCAATGTACAAAGAGCAAGCTGGGCTGGGCTGGAGACCTACC 3480
Db      ||| 5867 TTGCTCCCTCTGACCAATGTACAAAGAGCAAGCTGGGCTGGGCTGGAGACCTACC 5926
Qy      ||| 3481 CACATTAATTGACTACTGAAATTTCTCATGCGCAGGCTCGGATTTGACAAAGAACTGGAA 3540
Db      ||| 5927 CACATTAATTGACTACTGAAATTTCTCATGCGCAGGCTCGGATTTGACAAAGAACTGGAA 5986
Qy      ||| 3541 GCCTTCCAGATGCAATGAGAAAGCGCCAGAGATACAAAGACGGCGCTCATTTTGGGAT 3600
Db      ||| 5987 GCCTTCCAGATGCAATGAGAAAGCGCCAGAGATACAAAGACGGCGCTCATTTTGGGAT 6046
Qy      ||| 3601 CCGGATCTCGCTTCTCTACTACAGCTTCTTCAAGAGATTAGCAGCAAGTGCAGTCTCC 3660
Db      ||| 6047 CCGGATCTCGCTTCTCTACTACAGCTTCTTCAAGAGATTAGCAGCAAGTGCAGTCTCC 6106
Qy      ||| 3661 GCACCTGCTATTACAGACTCTGAGAACCGTATCGGCAAGTCGAGAAAGATATGCGAGG 3720
Db      ||| 6107 GCACCTGCTATTACAGACTCTGAGAACCGTATCGGCAAGTCGAGAAAGATATGCGAGG 6166
Qy      ||| 3721 TTGGTCAAAAAGAGAGATGAGAGACAGACAGGACCCCTACCTCGTCCCGCTCAACAG 3780
Db      ||| 6167 TTGGTCAAAAAGAGAGATGAGAGACAGACAGGACCCCTACCTCGTCCCGCTCAACAG 6226
Qy      ||| 3781 GTTTATGAAAAATGTCGCGCATCACGCTGAGGCGATGAGCAAAATCCGAGCAAAATTTAT 3840
Db      ||| 6227 GTTTATGAAAAATGTCGCGCATCACGCTGAGGCGATGAGCAAAATCCGAGCAAAATTTAT 6286
Qy      ||| 3841 GATTCTAAGGTATCAGGTGCTGAGACTGTCTTCTCGGAGCCGTTGAGATGATATCA 3900
Db      ||| 6287 GATTCTAAGGTATCAGGTGCTGAGACTGTCTTCTCGGAGCCGTTGAGATGATATCA 6346
Qy      ||| 3901 TTGGCATTTGCTGAGGSGCTAGACGGCTTTCAAGCTGTACTACCAAGAGACCCCAATTC 3960
Db      ||| 6347 TTGGCATTTGCTGAGGSGCTAGACGGCTTTCAAGCTGTACTACCAAGAGACCCCAATTC 6406
Qy      ||| 3961 GTGTGCGAGATGCGGCGCAGACAGCTCGCTACATTAAGGCGCAGATGACGAGACAGCC 4020
Db      ||| 6407 GTGTGCGAGATGCGGCGCAGACAGCTCGCTACATTAAGGCGCAGATGACGAGACAGCC 6466
Qy      ||| 4021 GGTGAAGCGCCCCCGGCTGTTGATGACCGCTTCAATGATGCGGCTTGTATGCGGATTAAG 4080
Db      ||| 6467 GGTGAAGCGCCCCCGGCTGTTGATGACCGCTTCAATGATGCGGCTTGTATGCGGATTAAG 6526
Qy      ||| 4081 AAGTTTACGAAGCAGTATGTGCGCAGGCTGAGGCGCATGATGCGAGTACCCCTGATCCG 4140
Db      ||| 6527 AAGTTTACGAAGCAGTATGTGCGCAGGCTGAGGCGCATGATGCGAGTACCCCTGATCCG 6586
Qy      ||| 4141 GAGGCTATGAAAGTGTGCGGCGATGATGTTTGTATGAAATGGTTTCAAGGGAATGGC 4200
Db      ||| 6587 GAGGCTATGAAAGTGTGCGGCGATGATGTTTGTATGAAATGGTTTCAAGGGAATGGC 6646
Qy      ||| 4201 GATTAT 4206
Db      ||| 6647 GATTAT 6652
```

. RESULT 3

```
CQ803434
LOCUS      CQ803434      4206 bp      DNA      linear      PAT 10-MAY-2004
DEFINITION Sequence 1 from Patent WO2004035784.
ACCESSION  CQ803434
VERSION    CQ803434.1  GI:47110295
KEYWORDS
SOURCE
ORGANISM   Neurospora crassa
            Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

REFERENCE
1          Mekeyev, E. and Bamford, D.
            Soluble rna polymerase protein and methods for the use thereof
            Patent: WO 2004035784-A 1 29-APR-2004;
            RNA-line Oy (FI)

FEATURES
            source
                1..4206
                /organism="Neurospora crassa"
                /mol_type="unassigned DNA"
                /db_xref="taxon:5141"

ORIGIN
Query Match      99.9%; Score 4202.8; DB 6; Length 4206;
Best Local Similarity 100.0%; Pred.No.0; Mismatches 2; Indels 0; Gaps 0;
Matches 4204; Conservative 0;

Qy      1  ATGAACTTATTACTCTAGAGAGAAATAGCCCCGTGAGGAATCATAAACCGGCTC 60
Db      1  ATGAACTTATTACTCTAGAGAGAAATAGCCCCGTGAGGAATCATAAACCGGCTC 60

Qy      61  AATTAAGACTTACAACCTGGGCTTCCAGTGTGTGCAACAACTTACCCCCCAGCC 120
Db      61  AATTAAGACTTACAACCTGGGCTTCCAGTGTGTGCAACAACTTACCCCCCAGCC 120

Qy      121  CGGAGAGCTGGCGGAGAGTGAAGGATTTTGGTGGCATAGCAAGATCTTACAGAGCC 180
Db      121  CGGAGAGCTGGCGGAGAGTGAAGGATTTTGGTGGCATAGCAAGATCTTACAGAGCC 180

Qy      181  CTGAACCTTCTACTATGCGGAGAGATGACTCCCTGTAACCAAGCAGAACCACTTCTTC 240
Db      181  CTGAACCTTCTACTATGCGGAGAGATGACTCCCTGTAACCAAGCAGAACCACTTCTTC 240

Qy      241  ATGAGGCCAAGACTGCGAGCTTGAACCTGGGTGCCAAAGCCCAAGCCGCTGACAG 300
Db      241  ATGAGGCCAAGACTGCGAGCTTGAACCTGGGTGCCAAAGCCCAAGCCGCTGACAG 300

Qy      241  ATGAGGCCAAGACTGCGAGCTTGAACCTGGGTGCCAAAGCCCAAGCCGCTGACAG 300
Db      241  ATGAGGCCAAGACTGCGAGCTTGAACCTGGGTGCCAAAGCCCAAGCCGCTGACAG 300

Qy      301  CTTCCGTTGCTCAAGGAACCTCCCGCGCGCTACTGCGCGCAACATGGGCAATTGGCAG 360
Db      301  CTTCCGTTGCTCAAGGAACCTCCCGCGCGCTACTGCGCGCAACATGGGCAATTGGCAG 360

Qy      361  ACTGTGTTGCTGAGGTGCTTAATAGTTTATGCCACTTCCCAATTAACACCAAGGTGA 420
Db      361  ACTGTGTTGCTGAGGTGCTTAATAGTTTATGCCACTTCCCAATTAACACCAAGGTGA 420

Qy      421  ACGTTTGGCAGAACTTAAGCGGCGCAAGTGGCTGAGCGGCCCAACCTCTACCAACAC 480
Db      421  ACGTTTGGCAGAACTTAAGCGGCGCAAGTGGCTGAGCGGCCCAACCTCTACCAACAC 480

Qy      481  AAAAGCAAGGATGAGCGCGCAATGTCACTTTGCTGATCCGCGCAAGGCTGTTGACT 540
Db      481  AAAAGCAAGGATGAGCGCGCAATGTCACTTTGCTGATCCGCGCAAGGCTGTTGACT 540

Qy      541  CGCTTTCGACAGAGTCTCTATTTCAGCGCGCGCGCATACCCCTAAAGTTTCCCGATCA 600
Db      541  CGCTTTCGACAGAGTCTCTATTTCAGCGCGCGCGCATACCCCTAAAGTTTCCCGATCA 600

Qy      601  GTGAATACCGGTTCCAAAGCAACCATCTTCGAGAGTGAAGATTCATATGATGACCAAG 660
Db      601  GTGAATACCGGTTCCAAAGCAACCATCTTCGAGAGTGAAGATTCATATGATGACCAAG 660

Qy      661  CGGCGCAAGGGCAAGGCTGTGATTAATGTTGCGGTGCGCGCGCGCGCGCGGCTATT 720
Db      661  CGGCGCAAGGGCAAGGCTGTGATTAATGTTGCGGTGCGCGCGCGCGCGCGGCTATT 720
```

QY 721 GCGAGCGCTTTGACAAAGTACCGACTCGAAGGCGCAATACAGAGATCCACGCGG 780  
| | | | |  
Db 721 GCGAGCGCTTTGACAAAGTACCGACTCGAAGGCGCAATACAGAGATCCACGCGG 780  
QY 781 ACAGGTATAGACGAGCGGACCAAGTGGATCTTTGATATCATCTCAAGGCACTTCTAT 840  
| | | | |  
Db 781 ACAGGTATAGACGAGCGGACCAAGTGGATCTTTGATATCATCTCAAGGCACTTCTAT 840  
QY 841 GGTTCAGAGTCTTTCAGCGCTTCCGCTCAATACAGACCTACCCAGAGATGTTTGGAG 900  
| | | | |  
Db 841 GGTTCAGAGTCTTTCAGCGCTTCCGCTCAATACAGACCTACCCAGAGATGTTTGGAG 900  
QY 901 GCTCTCTCTTCAAGCCGAGAGAGAGGCGCTGGATGCGCAGGCTTTGAGGCTGGA 960  
| | | | |  
Db 901 GCTCTCTCTTCAAGCCGAGAGAGAGGCGCTGGATGCGCAGGCTTTGAGGCTGGA 960  
QY 961 CACTTGATTGATCTCTAGCAAGAGAAACAACCAAGTCCCATATGATACCAAGCC 1020  
| | | | |  
Db 961 CACTTGATTGATCTCTAGCAAGAGAAACAACCAAGTCCCATATGATACCAAGCC 1020  
QY 1021 CTTTCAATGCTTCTCCAGGGTGAATCTTCTTCAAGCACTTATATGAGTGGTTTCAAGT 1080  
| | | | |  
Db 1021 CTTTCAATGCTTCTCCAGGGTGAATCTTCTTCAAGCACTTATATGAGTGGTTTCAAGT 1080  
QY 1081 TCCGCGGCGGAGGCGGCAATCCCGAGCGGATGCTCAATGGACTGGCTCCGAGGGA 1140  
| | | | |  
Db 1081 TCCGCGGCGGAGGCGGCAATCCCGAGCGGATGCTCAATGGACTGGCTCCGAGGGA 1140  
QY 1141 GAAAGCGCTCGATCTCAGGTTCAAGTTCAATGCCCCGGTGGTTGAGCTCGGCTAGAAAT 1200  
| | | | |  
Db 1141 GAAAGCGCTCGATCTCAGGTTCAAGTTCAATGCCCCGGTGGTTGAGCTCGGCTAGAAAT 1200  
QY 1201 ATTTGGCCGAATTTCCCAATGGCTTACAGAGCTCTCTCGCTTGGTCAATGGAAAT 1260  
| | | | |  
Db 1201 ATTTGGCCGAATTTCCCAATGGCTTACAGAGCTCTCTCGCTTGGTCAATGGAAAT 1260  
QY 1261 ACCGAGCTCTTATGCACTGCAAGTAGTGACTTGGAGCGAGAGCTGGGCTTAAAGTAC 1320  
| | | | |  
Db 1261 ACCGAGCTCTTATGCACTGCAAGTAGTGACTTGGAGCGAGAGCTGGGCTTAAAGTAC 1320  
QY 1321 GACCTTCTGCTGCTTACCGCGCGGATGTCACAGATATCTGGAAGACTCTTACCGGCT 1380  
| | | | |  
Db 1321 GACCTTCTGCTGCTTACCGCGCGGATGTCACAGATATCTGGAAGACTCTTACCGGCT 1380  
QY 1381 GATGCTTTCCGTGGTAAACCTTTTCCAGAAAAGCCGCCCAAGAGCTGTCGACCGCA 1440  
| | | | |  
Db 1381 GATGCTTTCCGTGGTAAACCTTTTCCAGAAAAGCCGCCCAAGAGCTGTCGACCGCA 1440  
QY 1441 ATGACGGGGCAACTTTGAGAGCAAGTAGTGCCTGCTCTGCTGCTTGAAGCTAC 1500  
| | | | |  
Db 1441 ATGACGGGGCAACTTTGAGAGCAAGTAGTGCCTGCTCTGCTGCTTGAAGCTAC 1500  
QY 1501 AATCCGACAACTGCGCTTACTGCGCCCTTACCTTGTGAGAGCTGAAGCCGCTATGTC 1560  
| | | | |  
Db 1501 AATCCGACAACTGCGCTTACTGCGCCCTTACCTTGTGAGAGCTGAAGCCGCTATGTC 1560  
QY 1561 GAGCAGGAGCTGTCAGCTCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
| | | | |  
Db 1561 GAGCAGGAGCTGTCAGCTCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
QY 1621 CCGTGGCTTACAGACCAAGCCCAAGTGTACCGCGGCTGTCAGCAAAACAACGAGTGG 1680  
| | | | |  
Db 1621 CCGTGGCTTACAGACCAAGCCCAAGTGTACCGCGGCTGTCAGCAAAACAACGAGTGG 1680  
QY 1681 GTCAAGAGATCATCAAGTGGCTCAAGTGGGCAACAATCTCTTGTAGAGCGCCCAATGG 1740  
| | | | |  
Db 1681 GTCAAGAGATCATCAAGTGGCTCAAGTGGGCAACAATCTCTTGTAGAGCGCCCAATGG 1740  
QY 1741 GCGCGCTTTCTTGGCAAAAGATGCGGATACAGAAAACCTCTCAAGGAGTTCCAGCTCGC 1800  
| | | | |  
Db 1741 GCGCGCTTTCTTGGCAAAAGATGCGGATACAGAAAACCTCTCAAGGAGTTCCAGCTCGC 1800

QY 1801 GCCGAGACCCGAAACCATCATCAAGAGAGATCCACTTCTTGGCCGAGACCGGCTT 1860  
| | | | |  
Db 1801 GCCGAGACCCGAAACCATCATCAAGAGAGATCCACTTCTTGGCCGAGACCGGCTT 1860  
QY 1861 ACCTTCCGACCTGATGTTCAAGAGAGATCTGTGCTTCCGCAAGAGAACTGTAGAG 1920  
| | | | |  
Db 1861 ACCTTCCGACCTGATGTTCAAGAGAGATCTGTGCTTCCGCAAGAGAACTGTAGAG 1920  
QY 1921 CAACGAGCCGAGTCAAAAGTATCAAAATGCTGAGACTGCTCTGCAACTGCAACAAC 1980  
| | | | |  
Db 1921 CAACGAGCCGAGTCAAAAGTATCAAAATGCTGAGACTGCTCTGCAACTGCAACAAC 1980  
QY 1981 ACTTGGACGCGACCTCAAGTGTCTCCGATCCAGCTCGGCTGATGAGACATAT 2040  
| | | | |  
Db 1981 ACTTGGACGCGACCTCAAGTGTCTCCGATCCAGCTCGGCTGATGAGACATAT 2040  
QY 2041 GCCATTATGACATTGGAGCTTCAACAGATCAGACACCAAGACCGATCTTCTTCACT 2100  
| | | | |  
Db 2041 GCCATTATGACATTGGAGCTTCAACAGATCAGACACCAAGACCGATCTTCTTCACT 2100  
QY 2101 TCAGGCACTGGGAGATGAAATGACCGGTGAGCCGCAATGTCGGAAGCTGGCCAG 2160  
| | | | |  
Db 2101 TCAGGCACTGGGAGATGAAATGACCGGTGAGCCGCAATGTCGGAAGCTGGCCAG 2160  
QY 2161 AGGATACGCGATGTTCTCGGTTGGGATGTCCTCTGCTGCAAGGGCGGTTGGT 2220  
| | | | |  
Db 2161 AGGATACGCGATGTTCTCGGTTGGGATGTCCTCTGCTGCAAGGGCGGTTGGT 2220  
QY 2221 TCGGCCAAGGAAATGGGGTTATCGAGTTGACGACACAGGCGATGAGATGGATCGAG 2280  
| | | | |  
Db 2221 TCGGCCAAGGAAATGGGGTTATCGAGTTGACGACACAGGCGATGAGATGGATCGAG 2280  
QY 2281 ACAATCCCGTCCAGCGCAAGTGGGAAATGCGACTTCTGTAATAAATCAACGTAACCTC 2340  
| | | | |  
Db 2281 ACAATCCCGTCCAGCGCAAGTGGGAAATGCGACTTCTGTAATAAATCAACGTAACCTC 2340  
QY 2341 GAAATCCGAGCGTGGCTTGAAGTCAAGTCAAGTCTCAACCTACAGCTGTTACT 2400  
| | | | |  
Db 2341 GAAATCCGAGCGTGGCTTGAAGTCAAGTCAAGTCTCAACCTACAGCTGTTACT 2400  
QY 2401 GTCTGGAGATGAGCGCAGGGAACAAGTGAAGATGCGCAAGCAATCGGTGACCGTCT 2460  
| | | | |  
Db 2401 GTCTGGAGATGAGCGCAGGGAACAAGTGAAGATGCGCAAGCAATCGGTGACCGTCT 2460  
QY 2461 ATCAAAGATTTGCAACGACGTTACGAGCAAAAGCATCTTGAATGCCCCAAGTGA 2520  
| | | | |  
Db 2461 ATCAAAGATTTGCAACGACGTTACGAGCAAAAGCATCTTGAATGCCCCAAGTGA 2520  
QY 2521 TTTGCGCAATGGGTTTACGAGAGTTATTCAGTTCGCGCAACTGAGTCAGCCAGCGCT 2580  
| | | | |  
Db 2521 TTTGCGCAATGGGTTTACGAGAGTTATTCAGTTCGCGCAACTGAGTCAGCCAGCGCT 2580  
QY 2581 GTGCTTTTCTTGTGGGCTTACCTGACAGTCAAGAGGACACTGAACTTCTTGAATGAC 2640  
| | | | |  
Db 2581 GTGCTTTTCTTGTGGGCTTACCTGACAGTCAAGAGGACACTGAACTTCTTGAATGAC 2640  
QY 2641 AGTGGTTCATCCCAAGAGCAAAAGTACTTGCAGACATCCGCTGGGATCTTCAAAAG 2700  
| | | | |  
Db 2641 AGTGGTTCATCCCAAGAGCAAAAGTACTTGCAGACATCCGCTGGGATCTTCAAAAG 2700  
QY 2701 CGGAAATGTGACAGTGAAGTCAAGCTGAACATCCGCTGCTGATGAGATACAT 2760  
| | | | |  
Db 2701 CGGAAATGTGACAGTGAAGTCAAGCTGAACATCCGCTGCTGATGAGATACAT 2760  
QY 2761 TACATGATTCGCAATTTCTGGGGTGTCTTGAAGAAATAGAGTTCTGCTCC 2820  
| | | | |  
Db 2761 TACATGATTCGCAATTTCTGGGGTGTCTTGAAGAAATAGAGTTCTGCTCC 2820  
QY 2821 TCAAGTTCAGACAGAGAGAGTCTTTTACACTCTTATCGAGCTGTGATGCTCTG 2880  
| | | | |  
Db 2821 TCAAGTTCAGACAGAGAGAGTCTTTTACACTCTTATCGAGCTGTGATGCTCTG 2880  
QY 2881 GCGGATGCCAGCCCATTTCCCTAGTATCCAAAGGGTTGAGAGACTTCAAGCA 2940  
| | | | |

Db 2881 GCGGATCCCGAGCCATTTCCCTAGATATCCAGGGGTGAGAGCTTTCAGCCCA 2940  
Qy 2941 GAGCTCACAGTCTCAAGATGTATATCATCTTCTCTAAGAGATGTACGGCTTGT 3000  
Db 2941 GAGCTCACAGTCTCAAGATGTATATCATCTTCTCTAAGAGATGTACGGCTTGT 3000  
Qy 3001 AAGAGCTATCTGGTGGAGACTAGCAGCGCGATATGGCCTTGGTCTGGTGGATCCGAG 3060  
Db 3001 AAGAGCTATCTGGTGGAGACTAGCAGCGCGATATGGCCTTGGTCTGGTGGATCCGAG 3060  
Qy 3061 ATGTGATGTGTTTCTGATATGGGAAATGCTCTGAGAGCCGACCTGTCTAGTACTTA 3120  
Db 3061 ATGTGATGTGTTTCTGATATGGGAAATGCTCTGAGAGCCGACCTGTCTAGTACTTA 3120  
Qy 3121 AAGAGGACAAAGACTTTCAAACACTTATGGCCTCAGACGGGCTCAGCGGCC 3180  
Db 3121 AAGAGGACAAAGACTTTCAAACACTTATGGCCTCAGACGGGCTCAGCGGCC 3180  
Qy 3181 AAGAGGACACTATCATATGATATGATCAGAGAGCTTCCATTTCCGCTCAGGCCAAC 3240  
Db 3181 AAGAGGACACTATCATATGATATGATCAGAGAGCTTCCATTTCCGCTCAGGCCAAC 3240  
Qy 3241 TTCTTGGGATGTGCACTAATCAAGAAAGGCTCTGTACATCAACATAGTGTCT 3300  
Db 3241 TTCTTGGGATGTGCACTAATCAAGAAAGGCTCTGTACATCAACATAGTGTCT 3300  
Qy 3301 AAGAGCGGACCATATCTTCTTAACTGCTGAGGAAACCTCTGATCAGAGCAAGCAA 3360  
Db 3301 AAGAGCGGACCATATCTTCTTAACTGCTGAGGAAACCTCTGATCAGAGCAAGCAA 3360  
Qy 3361 GGTATGTCTTTTAAAGAGCAAGCTGGGCTCAATGGGTAGGGAACCTGTTGGCGGTGA 3420  
Db 3361 GGTATGTCTTTTAAAGAGCAAGCTGGGCTCAATGGGTAGGGAACCTGTTGGCGGTGA 3420  
Qy 3421 TTGTCCCTTCTTACCCCAATGTACAGAGGACAGTGGTCCGGCGCGGAGAGCCCTACC 3480  
Db 3421 TTGTCCCTTCTTACCCCAATGTACAGAGGACAGTGGTCCGGCGCGGAGAGCCCTACC 3480  
Qy 3481 CACATTATGACTACCTGAATTTCTCCATCGCAGGCTCGATTCAGAACAGAACTGGA 3540  
Db 3481 CACATTATGACTACCTGAATTTCTCCATCGCAGGCTCGATTCAGAACAGAACTGGA 3540  
Qy 3541 GCGCTTCCACATGCCATGAAGCGCGCCAGAGATACAGAGACGGCGCTCACTTTGGAT 3600  
Db 3541 GCGCTTCCACATGCCATGAAGCGCGCCAGAGATACAGAGACGGCGCTCACTTTGGAT 3600  
Qy 3601 CCGGATCTCGCTTCTTACCTACCTTCTTCAAGAGATTAAGCAAGTGGCATGCTCC 3660  
Db 3601 CCGGATCTCGCTTCTTACCTACCTTCTTCAAGAGATTAAGCAAGTGGCATGCTCC 3660  
Qy 3661 GCACTGTATTCAGACTCTGAAGAACCTGATCGCGAAGTCCAGAAAGATATGCGAG 3720  
Db 3661 GCACTGTATTCAGACTCTGAAGAACCTGATCGCGAAGTCCAGAAAGATATGCGAG 3720  
Qy 3721 TTGGTCAAAAACAGAGATGAGAGACAGAGACCCCTTACCTGTGCGCGTCAACAG 3780  
Db 3721 TTGGTCAAAAACAGAGATGAGAGACAGAGACCCCTTACCTGTGCGCGTCAACAG 3780  
Qy 3781 GTTTATGAAAATGAGTGGCGCATCAGCGCTGAGCGCATGGAACAATCCGAGCAAAATTAT 3840  
Db 3781 GTTTATGAAAATGAGTGGCGCATCAGCGCTGAGCGCATGGAACAATCCGAGCAAAATTAT 3840  
Qy 3841 GATTCTAAGGTATCAGGTGCTGAGAGTCTCTTCTCGCGACCGTGAATGATATCA 3900  
Db 3841 GATTCTAAGGTATCAGGTGCTGAGAGTCTCTTCTCGCGACCGTGAATGATATCA 3900  
Qy 3901 TGGGCAATGTCTGAGGCTAGACGGCTTTCAGAGTGTATCAACAAGAGCCCAAGTTTC 3960  
Db 3901 TGGGCAATGTCTGAGGCTAGACGGCTTTCAGAGTGTATCAACAAGAGCCCAAGTTTC 3960  
Qy 3961 GTTGTGAGATGGCGGACAGAGCTCGCTTACCTTAAGCGGCAAGTGAAGACCAACCC 4020  
Db 3961 GTTGTGAGATGGCGGACAGAGCTCGCTTACCTTAAGCGGCAAGTGAAGACCAACCC 4020

Db 3961 GTTGTGAGATGGCGGACAGAGCTCGCTTACCTTAAGCGGCAAGTGAAGACCAACCC 4020  
Qy 4021 GGTGAAGGCGCCCGGCGTTGATGACCGGCTTATGATGCGGGCTTGAATGCGGATTAAG 4080  
Db 4021 GGTGAAGGCGCCCGGCGTTGATGACCGGCTTATGATGCGGGCTTGAATGCGGATTAAG 4080  
Qy 4081 AAGTTTACGAGAGATGTGCGCAGGCTGAGAGGCGGATGATCGAGATACCTGATCCG 4140  
Db 4081 AAGTTTACGAGAGATGTGCGCAGGCTGAGAGGCGGATGATCGAGATACCTGATCCG 4140  
Qy 4141 GAGCTTATGAATGCTGCGGCGATGATGATTTTGAATGAATTTGTTCAAGGAAATGGC 4200  
Db 4141 GAGCTTATGAATGCTGCGGCGATGATGATTTTGAATGAATTTGTTCAAGGAAATGGC 4200  
Qy 4201 GATTAT 4206  
Db 4201 GATTAT 4206

RESULT 4  
CO803436 3078 bp DNA linear PAT 10-MAY-2004  
LOCUS Sequence 3 from Patent WO2004035784.  
DEFINITION CO803436  
ACCESSION CO803436 GI:47110296  
VERSION  
KEYWORDS  
SOURCE Neurospora crassa  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetiales; Sordariaceae; Neurospora.

REFERENCE 1  
AUTHORS Mekeyev, E. and Bamford, D.  
TITLE Soluble rna polymerase protein and methods for the use thereof  
JOURNAL Patent: WO 2004035784-A 3 29-APR-2004;  
RNA-line Oy (FI)

FEATURES  
Source Location/Qualifiers  
1..3078 /organism="Neurospora crassa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5141"

ORIGIN  
Query Match 73.2%; Score 3078; DB 6; Length 3078;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3078; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1129 GCTCGAGGAGAAAGCGCTGATCTCAGTTCAAGTTCAGGCCCGGTTGGTAGCT 1188  
Db 1 GCTCGAGGAGAAAGCGCTGATCTCAGTTCAAGTTCAGGCCCGGTTGGTAGCT 60  
Qy 1189 CGGCTGAGAAATTTTGGCCGAAATTTCCCAATGAGCTACAGAAAGCTCTCGCTGT 1248  
Db 61 CGGCTGAGAAATTTTGGCCGAAATTTCCCAATGAGCTACAGAAAGCTCTCGCTGT 120  
Qy 1249 GCATGGGAAGTTACCAAGACTCTTTATGCACTGCAAAAGTATGACTTGAAGACGAGCTTG 1308  
Db 121 GCATGGGAAGTTACCAAGACTCTTTATGCACTGCAAAAGTATGACTTGAAGACGAGCTTG 180  
Qy 1309 GGCCTAAGTACGAGACCTTCTCGGTCTACCGGCGGAGTACACAGATATCTGGAAGACT 1368  
Db 181 GGCCTAAGTACGAGACCTTCTCGGTCTACCGGCGGAGTACACAGATATCTGGAAGACT 240  
Qy 1369 CTCTACCGGCTTGATGCTTTCCGTGTAAACCTTTCCAGAAAAGCGGCCCAACGACGTG 1428  
Db 241 CTCTACCGGCTTGATGCTTTCCGTGTAAACCTTTCCAGAAAAGCGGCCCAACGACGTG 300  
Qy 1429 TTGCTGACCGCAATGACGGGCAACTTTGAGAGCAAAAGTATGCGCTTGTCTCTGCT 1488  
Db 301 TTGCTGACCGCAATGACGGGCAACTTTGAGAGCAAAAGTATGCGCTTGTCTCTGCT 360  
Qy 1489 GTTCTAGACTACCAATCGGACAACTGCGCTACCTGCGCCCTTACCTTGAAGCTGAAG 1548  
Db 361 GTTCTAGACTACCAATCGGACAACTGCGCTACCTGCGCCCTTACCTTGAAGCTGAAG 420



QY 1549 CCGCTCATGTTGAGCAGAGCGTGTGACCTCACCCGTCGGTTCGCTGATAGGTTTTC 1608  
DB 421 CCGCTCATGTTGAGCAGAGCGTGTGACCTCACCCGTCGGTTCGCTGATAGGTTTTC 480  
QY 1609 GAGATCTTATACCGTTCGCTACAGCACAAGCCCAAGTGTACCGCGGTGTACAGCAA 1668  
DB 481 GAGATCTTATACCGTTCGCTACAGCACAAGCCCAAGTGTACCGCGGTGTACAGCAA 540  
QY 1669 CAACCAAGTGTGCTGAGAAAGTCACTCAGTGTCTCAGATGGGCAACATTTCTTTGTA 1728  
DB 541 CAACCAAGTGTGCTGAGAAAGTCACTCAGTGTCTCAGATGGGCAACATTTCTTTGTA 600  
QY 1729 GCGCGCCAAATGCGCGCTTTCTTTCGCAAGATGCGGATACAGGAAACCTCTCAGGAG 1788  
DB 601 GCGCGCCAAATGCGCGCTTTCTTTCGCAAGATGCGGATACAGGAAACCTCTCAGGAG 660  
QY 1789 TTCCAGCTCCGCGCGGAGAACCCCAATCATCAGAGAGAGTCACTTCTTTC 1848  
DB 661 TTCCAGCTCCGCGCGGAGAACCCCAATCATCAGAGAGAGTCACTTCTTTC 720  
QY 1849 GAGACCGGATTAAGTCTTCCGACCTGATGTGTTCAAGAGAGATCTGTCTTCCGACAG 1908  
DB 721 GAGACCGGATTAAGTCTTCCGACCTGATGTGTTCAAGAGAGATCTGTCTTCCGACAG 780  
QY 1909 GAACCTGTAGAGCAACGAGCCGAGTTCAAAGTTAGTCAAATGTGCTGAGTCTGCTTCCGA 1968  
DB 781 GAACCTGTAGAGCAACGAGCCGAGTTCAAAGTTAGTCAAATGTGCTGAGTCTGCTTCCGA 840  
QY 1969 CTGCAACAACAACCTTGGCAGCCGACCTCAAGTTGTTCTCCGATACAGCTCGGTCTG 2028  
DB 841 CTGCAACAACAACCTTGGCAGCCGACCTCAAGTTGTTCTCCGATACAGCTCGGTCTG 900  
QY 2029 AGTAAGACATATGCTATATGACATTTGAGGCTCACGATACACACACAACAGCCAT 2088  
DB 901 AGTAAGACATATGCTATATGACATTTGAGGCTCACGATACACACACAACAGCCAT 960  
QY 2089 CTTCTTCACTTCGAGCACTGGCGAAGTGAATGAAATGACGCTGTAGGCGCATGCGGA 2148  
DB 961 CTTCTTCACTTCGAGCACTGGCGAAGTGAATGAAATGACGCTGTAGGCGCATGCGGA 1020  
QY 2149 AGCGTGGCAAGAGATACCGGATGTTCTCGGTTTGGTGTATGTGCTCTGTGTGCA 2208  
DB 1021 AGCGTGGCAAGAGATACCGGATGTTCTCGGTTTGGTGTATGTGCTCTGTGTGCA 1080  
QY 2209 GGGCGGTTGGTGTGGCGCAAGGGAATGTGGTTATCGACGTTGACGACAGCGCATGAG 2268  
DB 1081 GGGCGGTTGGTGTGGCGCAAGGGAATGTGGTTATCGACGTTGACGACAGCGCATGAG 1140  
QY 2269 GATTGATCGAGACATACCCGTCGAGCGCAAGTGGGAATGCACTTGTGATTAACAT 2328  
DB 1141 GATTGATCGAGACATACCCGTCGAGCGCAAGTGGGAATGCACTTGTGATTAACAT 1200  
QY 2329 CAACGTAACCTCGAAGTCGAGCGGTCTTCTGAATCGAAGTCAAGTGTCTCAACCTA 2388  
DB 1201 CAACGTAACCTCGAAGTCGAGCGGTCTTCTGAATCGAAGTCAAGTGTCTCAACCTA 1260  
QY 2389 CAGCTGTAACTGTCTTGAAGATGAGCCGAGCAAGGTGAAGATGCGCAGCAATC 2448  
DB 1261 CAGCTGTAACTGTCTTGAAGATGAGCCGAGCAAGGTGAAGATGCGCAGCAATC 1320  
QY 2449 GGTGACGCTTATCAACGATTTGCAACGACGATTCAGCGAGCAAAAGCATGTTGAT 2508  
DB 1321 GGTGACGCTTATCAACGATTTGCAACGACGATTCAGCGAGCAAAAGCATGTTGAT 1380  
QY 2509 CGCCAGTGAATTTCCGCAATGGGTTTACGAGATTAATTCAGTCCGCAACTCGATC 2568  
DB 1381 CGCCAGTGAATTTCCGCAATGGGTTTACGAGATTAATTCAGTCCGCAACTCGATC 1440  
QY 2569 AGCCACGGCGGTGTCTTTTCTGTGCTGGGCTACTGACAGTCAAGAGAGACACTGAAC 2628  
DB 1441 AGCCACGGCGGTGTCTTTTCTGTGCTGGGCTACTGACAGTCAAGAGAGACACTGAAC 1500

QY 2629 TTCTTGATGAACAGTGGTTCGATCCCAAGAGCAAAAGTACTTGCAAGACATCGCTCG 2688  
DB 1501 TTCTTGATGAACAGTGGTTCGATCCCAAGAGCAAAAGTACTTGCAAGACATCGCTCG 1560  
QY 2689 GATCTTCAAAAACGGAATGTGACAGCTTGAAGTCCAACTGAAATCTCCGTGCTGCA 2748  
DB 1561 GATCTTCAAAAACGGAATGTGACAGCTTGAAGTCCAACTGAAATCTCCGTGCTGCA 1620  
QY 2749 TCACATATCATTTATCATGATTTCTGCGGTGTGCTTGAAGAAAAGAGTTTAT 2808  
DB 1621 TCACATATCATTTATCATGATTTCTGCGGTGTGCTTGAAGAAAAGAGTTTAT 1660  
QY 2809 GTCCGATTTCTCCAAAGTTTCAGAGACAGAGAGAGTCTTTTACATCTCTACGACTGT 2868  
DB 1681 GTCCGATTTCTCCAAAGTTTCAGAGACAGAGAGAGTCTTTTACATCTCTACGACTGT 1740  
QY 2869 GATGTCTCTGTGGCGATCCCAAGCCCATTTCCCTAGATATCCAAAGGTTTCAGACA 2928  
DB 1741 GATGTCTCTGTGGCGATCCCAAGCCCATTTCCCTAGATATCCAAAGGTTTCAGACA 1800  
QY 2929 GTCTTCAAGCCAGCTCCACAGTCTCAAGATGTATCATCTTCTCTACTTAAAGAGAT 2988  
DB 1801 GTCTTCAAGCCAGCTCCACAGTCTCAAGATGTATCATCTTCTCTACTTAAAGAGAT 1860  
QY 2989 GTACCGCTTGCTAAGAGCTATCTGTGAGACTACAGACGCGATATGCGCTGGGTCTGC 3048  
DB 1861 GTACCGCTTGCTAAGAGCTATCTGTGAGACTACAGACGCGATATGCGCTGGGTCTGC 1920  
QY 3049 TGGGATCCGAGATCGTGTGATGTTTCTGCAATGCGGAATGCTCTGAGCCGACCTG 3108  
DB 1921 TGGGATCCGAGATCGTGTGATGTTTCTGCAATGCGGAATGCTCTGAGCCGACCTG 1980  
QY 3109 TCTAGTACCTTAAAGAGAGCAAAACGATTTCAAAACATTAATGCGCTCACAGGACG 3168  
DB 1981 TCTAGTACCTTAAAGAGAGCAAAACGATTTCAAAACATTAATGCGCTCACAGGACG 2040  
QY 3169 GGCCTCAGCGGCAAGAGACAGACTACATACGATGTGATCCAGAGAGCTTCCATTTGCC 3228  
DB 2041 GGCCTCAGCGGCAAGAGACAGACTACATACGATGTGATCCAGAGAGCTTCCATTTGCC 2100  
QY 3229 CTGAGCCCAACTTCTTGGGATGTGACTTAACCTAACAAAGAGCTCTGTATCATCAAC 3288  
DB 2101 CTGAGCCCAACTTCTTGGGATGTGACTTAACCTAACAAAGAGCTCTGTATCATCAAC 2160  
QY 3289 AATAGTGTCTTAAACAGCCGCGCATATCTTATGTTCACTGTGTGGGAAACTGTGAT 3348  
DB 2161 AATAGTGTCTTAAACAGCCGCGCATATCTTATGTTCACTGTGTGGGAAACTGTGAT 2220  
QY 3349 CAGAGCAAGCAAGGATTTGTCTTAAAGAGCAAGTGGGCTCAATGTGCGTAAGGAACTG 3408  
DB 2221 CAGAGCAAGCAAGGATTTGTCTTAAAGAGCAAGTGGGCTCAATGTGCGTAAGGAACTG 2280  
QY 3409 CTTGGCGGTGATTTGCTCTTCTGACCCAAATGTACAAAGACGACAGTTGGCTGGGCGC 3468  
DB 2281 CTTGGCGGTGATTTGCTCTTCTGACCCAAATGTACAAAGACGACAGTTGGCTGGGCGC 2340  
QY 3469 GGAAGGCTACCCCATTTATGATTAACCTGAAATTTCTCAATGCGGACGCTGCGATGAC 3528  
DB 2341 GGAAGGCTACCCCATTTATGATTAACCTGAAATTTCTCAATGCGGACGCTGCGATGAC 2400  
QY 3529 AAGGAATGGAAGGCTTCCAAATGTCATGAAACGCGCAAGATACAGAAAGCGGCGCT 3588  
DB 2401 AAGGAATGGAAGGCTTCCAAATGTCATGAAACGCGCAAGATACAGAAAGCGGCGCT 2460  
QY 3589 CACTTTGGGATCCGATCTTCTCTTCTCAACGTTCTTCAAGAGATTAAGGACAG 3648  
DB 2461 CACTTTGGGATCCGATCTTCTCTTCTCAACGTTCTTCAAGAGATTAAGGACAG 2520  
QY 3649 TCGGATCTGCTCCGACCTGCTATTCAGACTCTGAAGAACCTATTCGCGGAAGTGAAGAA 3708  
DB 2521 TCGGATCTGCTCCGACCTGCTATTCAGACTCTGAAGAACCTATTCGCGGAAGTGAAGAA 2560  
QY 3709 GAATATGAGAGTTGTCAAAAACAAAGAGATGAGACAGCAAGACCCCTACCTGTCT 3768

Db 2581 GAATATGGCAGGTTGTCAAAAACAAAGAGATGAGACAGCAAGCAACCCCTACCCCTGTC 2640  
QY CGCGTCAACCAAGGTTTATGAAAAATGTTGCGCCATCAAGCCTGAGGCGATGACAAATCC 3828  
Db 2641 CGCGTCAACCAAGGTTTATGAAAAATGTTGCGCCATCAAGCCTGAGGCGATGACAAATCC 2700  
QY 3829 GAGCAAAATATGATCTTAAAGTATCAGGTTGCTGAGAGCTGTCTCTCGCGAACCGT 3888  
Db 2701 GAGCAAAATATGATCTTAAAGTATCAGGTTGCTGAGAGCTGTCTCTCGCGAACCGT 2760  
QY 3889 GAGTAAATCATGGGATGCTGAGGGCTAGCAAGCCTTCAAGCTGTACTACCAAG 3948  
Db 2761 GAGTAAATCATGGGATGCTGAGGGCTAGCAAGCCTTCAAGCTGTACTACCAAG 2820  
QY 3949 AGCCCCAAGTTCGTGTGCGAGATGCGGCGAGACAGCTCGCGTACATTAAAGCGCAGATG 4008  
Db 2821 AGCCCCAAGTTCGTGTGCGAGATGCGGCGAGACAGCTCGCGTACATTAAAGCGCAGATG 2880  
QY 4009 ACGAGCAGACCCGGTGAAGCGCCCGCGCTTGATGACCCGCTTCATGTATCGCGCTTG 4068  
Db 2881 ACGAGCAGACCCGGTGAAGCGCCCGCGCTTGATGACCCGCTTCATGTATCGCGCTTG 2940  
QY 4069 ATGCGGATTAAGAGTTTACGAAGCAGTATGTGSCAGCGCTGAGAGGCGCATGATCGGAG 4128  
Db 2941 ATGCGGATTAAGAGTTTACGAAGCAGTATGTGSCAGCGCTGAGAGGCGCATGATCGGAG 3000  
QY 4129 TACCGTATCCGAGGCTCTATGAGTCTGTGCGCATGATGATTTTGAATGATGTTTC 4188  
Db 3001 TACCGTATCCGAGGCTCTATGAGTCTGTGCGCATGATGATTTTGAATGATGTTTC 3060  
QY 4189 ACAGGGAATGCGCATTTAT 4206  
Db 3061 ACAGGGAATGCGCATTTAT 3078  
RESULT 5  
AF429315 125020 bp DNA linear PRI 18-JAN-2002  
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.  
ACCESSION AF429315  
VERSION AF429315.1 GI:17646244  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 125020)  
AUTHORS Holmes S.E., O'Hearn E., Rosenblatt A., Callahan C., Hwang H.S.,  
Ingber S., Ashworth R.G., Fleisher A., Stevanin G., Brice A.,  
Potter N.T., Ross C.A. and Margolis R.L.  
A repeat expansion in the gene encoding junctophilin-3 is  
associated with Huntington disease-like 2  
Nat. Genet. 29 (4), 377-378 (2001)  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
Source  
1. 125020  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16q24.3; Between D16S520 and WI-12410"  
/note="Isolated from a patient with Huntington's  
Disease-Like 2 (HDL2)"  
complement(35581..35746)  
/rpt\_type=tandem  
/rpt\_unit="ctg"

gene complement(<36507..>36887)  
/gene="JPH3"  
/note="synonym: JP3"  
mRNA complement(<36507..>36887)  
/gene="JPH3"  
/product="Junctophilin 3"  
complement(<36507..36887)  
/gene="JPH3"  
/note="component of the junctional complex between plasma  
membrane and endoplasmic reticulum"  
/codon\_start=1  
/product="Junctophilin 3"  
/protein\_id="AA140941.1"  
/db\_xref="GI:17646245"  
/translacion="MSSGRFNPDDGSGYCGGMDKRAHGVCTGPKGGEYTGMS  
HGEFVGVYTWPSGNTYOGTAAOGKNGHIGLESKKMYKGEWTFGKRGVRECA  
NGAKYEGTMSNGLDGYETETSDG"  
ORIGIN  
Query Match 1.3%; Score 55.2; DB 9; Length 125020;  
Best Local Similarity 10.6%; Pred. No. 0.01;  
Matches 102; Conservative 397; Mismatches 460; Indels 2; Gaps 1;  
QY 1389 CCCTGTGAACCCCTTCCAGAAAAACCCGCCAACAGAGCTGTGTAACCGCAATGACGG 1448  
Db 16887 MCYSSKMSMBMSVSVYVMSHMASBSQMBKCMTMSCCMSMYKSSWGSWMGCC 16946  
QY 1449 CAATTTGAGAGAGAAAGTATGAGTCCCGTCTCTCTGCTGTCTAGACTAATCCGA 1508  
Db 16947 WGRRRSKMKWYRGRMSRSMYRGTGSKMRMSNMCTSSCYVSMCMCWSGCCMRSCC 17006  
QY 1509 CAATCGCTTACTGCGCCCTTTACTTGTGAAGCTGAAGCCGCTCATGTTGACGAGG 1568  
Db 17007 CMCRSYCCMRVCCACACCYMSYWTMSASYSMSRYSKMSKMSRSMCRSRK 17066  
QY 1569 CTGTGACTACCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1628  
Db 17067 CGSWMGKGGKSGYGRKTKRKSNGKMGAKMYRYSRSMKMYSSKMYCMYCMWGR 17126  
QY 1629 TACGAGCAGCAGCCCAAGGTATACCGCGGTGTGACAAACAACAGGTGCGTCGAAGA 1688  
Db 17127 RGCYCGMTSRSMCMSCSYCAKAKSMCYSYVMSMKGYVYASVYSGMSKXCMRGSTY 17186  
QY 1689 AGTCATCTCACTGCTCAACGATGGGCAACATTTCTTTGAGCCGCCAATGCGCGCTT 1748  
Db 17187 STSCGCGCTTTTCCCCCNANNTGGGAAGCTTTTNCNKYRYSYRRKNGCMKYNVYN 17246  
QY 1749 CTTCGCCAAAGATCGCGGATACAGAAACCTCTCAGAGAGTCCAGTCCGCGCGAGA 1808  
Db 17247 NSWRSSCRAMSCCTKTKSNTMSMASYCMWMSMYCSMRSMASRMSWSYMMKMSWR 17306  
QY 1809 CCGAAACCCATCATGAAGAGAG--AGTCCATCTCTTGGCCGAGACCGCATTAAGTTTC 1866  
Db 17307 MSYVMSKCMKMKYCMMSMRSGMSYVYVMSKSSRGTCTRCYCMWSSSCVSKSY 17366  
QY 1867 CGACTGATGTCTTAAAGCAGATGCTGTCTTCCGCGAGAGAACCTTAGAGCAACG 1926  
Db 17367 MMARSKMKMKMSRSMWGSWMSRSCVCKYKMSRSMSCVRCACGMMMKGGYMR 17426  
QY 1927 ACGAGTTCAAGTTAGTAAAGTCGAGCTGCTCTCTGTAACGTGACAAACAACACTTG 1986  
Db 17427 YMRCSMKRWMKSMYRMRWRKRGAMMMCMKYSRSMCMKSYCASCRSCAMM 17486  
QY 1987 CAGCCGACCTCAAGTTGTTCTCCGTATCCAGCTCGTCTGATGAATATGCGATT 2046  
Db 17487 SGTVMCASVMCMSSYMTCTWCTWSYVYRCTCWCMSYTSKSMSSMSYKXGRK 17546  
QY 2047 ATGACATTTGAGGCTTACAGATCAGACACACAGAACCGATTTCTTACCTTAGGC 2106  
Db 17547 SYMCTSRGAMSCRRCCYMRGASMBRAGSMRRAKRSMGGRKMYMTGGMRSXY 17606  
QY 2107 ACTGGGAAGTATGATGACCGTGTAGCCCGCATGTCGGAAGCGTGGCGAAGAGATA 2166



Db 17607 YCTGRHMMTVMCCWRBRSSMYRRSMAWGMKRSWSGMRMMGASASRRCKGASRSSW 17666  
Qy 2167 CGCATGTTCTCGGTTGGTGATGTCCTCTCTGCAAGGCGGTTGGTCCGCC 2226  
Db 17667 CSRMKMRGSCWSSKMGSSRSASCKSGMRMRSSSKYRRGKRSMTKS 17726  
Qy 2227 AAGGAATGTCGTTATGACGCTTGACACACACGCGATGATGATGATGACATAC 2286  
Db 17727 KGSAGKSKMRSSMTSSCTYYASSCMMMSKSCMCCMMKRCACCTSSMCTSM 17786  
Qy 2287 CCGTCCAGCGCAAGTGGAATGCGACTTCGTATTAACATCAACGTACCTCGAATC 2346  
Db 17787 YRCMGWMSYSTGCTTCKCTGKSRMTWMMYMTSKMYMRAACMYCYMSTRMM 17846  
Qy 2347 C 2347  
Db 17847 C 17847

RESULT 6  
AF468822 6688 bp DNA linear PLN 14-FEB-2002  
LOCUS Diaporthe perijuncta putative RNA-dependent RNA polymerase RDP-1  
AF468822  
AF468822.1 GI:18656896

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Diaporthe perijuncta  
Diaporthe perijuncta  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Diaporthales; Valsaceae; Diaporthe.  
1 (bases 1 to 6688)  
Preisig, O. and Wingfield, M.J.  
Putative fungal RNA-dependent RNA polymerase  
Unpublished  
2 (bases 1 to 6688)  
Preisig, O. and Wingfield, M.J.  
Direct Submission  
Submitted (15-JUN-2002) Forestry and Agricultural Biotechnology  
Institute, University of Pretoria, Lunnon Street 74, Pretoria 0002,  
South Africa

FEATURES  
source  
Location/Qualifiers  
1..6688  
/organism="Diaporthe perijuncta"  
/mol\_type="genomic DNA"  
/isolate="CMW3407"  
/db\_xref="taxon:186170"  
/country="South Africa"  
join(<1..275,348..>381)  
/product="unknown"  
join(<1..275,348..381)  
/note="similar to D-lactate dehydrogenase (D-lactate  
fericytochrome C oxidoreductase)"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAL78033.1"  
/db\_xref="GI:18656897"  
/translation="MDDGLFASILGHIGDNFHESLIYNFNGEREKVECVKMYR  
RAIMDGTGGEHSIGMKDLSLLEVGPELTAVMRAIKSLDPRMIMNPKIPDRS  
"  
1361..>5870  
/gene="rdp-1"  
join(<1361..4598,4654..>5870)  
/gene="rdp-1"  
/product="putative RNA-dependent RNA polymerase RDP-1"  
join(1361..4598,4654..5870)  
/gene="rdp-1"  
/codon\_start=1  
/product="putative RNA-dependent RNA polymerase RDP-1"  
/protein\_id="AAL78034.1"  
/db\_xref="GI:18656898"  
/translation="MSHHSPTPTPDGSAQPDASAQIHHTTQTSIPDATAKDNPR  
ISTLRPAIADTSDASNRNTPSVRLPPLPLSMSTHSNGNVLGQKVSLSGPIQGRS

IOINARTFEGPHDPNPRONRPNRROORSPOSILQORSTPAPLTGAVALSAQSLTI  
DLFLPTDIKIREITWESMTSNLNVNRVEIYENPDGKRGKAVIIEPPIRMPYMS  
PMLITIDPADSSLOVLETPEPFPSTETPCNRPI PAPOCKKILPTKLVFPMMS  
SAILNMPORSYEDVILLTRSKRQFELQTVNIGRGCAFKSQIDPSHTKIXIKGR  
ENAMVLAAILPSPPLVGRLPVSHDEDTLMOMELMARQTISSBMLRLPALU  
EKQWTSLIEGRMTTYLLDMESLSMESRWNTLADLVNKLILQVHDFLNPGLSNM  
KLIDNASEMTLELIGSTQYHLP PVRVQLEVCISGGLSEYVNAQLEKLSIFED  
RALLMEGVAENAKOFFEPPEMFDRIILHYVNAIIPYATLVRAVITPTIYFET  
PCVELNRLIRKYSDLNDRRLVQPTDEYFCKI PESSQNSKDNLTTRVRYMONG  
IIGDRHYKTLAPNSOPRENGAFPFETDHTVTCASIRNMKGRFRHRSVCKPAAKMO  
CPTTQVNGISIPKIRIGDIQRSGDNWNTFDVGKISVFPAHILASERLPEP  
SCPOMMGCKGKVVMVDPDI PANEVHIRSOEKFVAVNGLEIKSTFHAHLNMQV  
IPVLIALGVDAVFMVMDLDELKEYEDALSMKASELRSQVDEQNTITVMENVD  
FMDSBPMTTLRLMKCWILKELKOKFALSVGSAIMI PCVDETGLVRSHQSDTEK  
GPNRISLPOIPLQVPIEGSDKSTNVEYITGICVGNRPALHPDVRVVAVDPE  
LRLHKVVPFPKTDGDRVPSMCGGDMGDVYVYDDELIPSEMHPRPLDHDADST  
PGIDSPADVTNGDVTFFAOMYNDLSLRIATLAFADQDLAGVHAKPKEIELAKHS  
KAVDYIKSGPAPVMEKHLQPRMNMHMERKSSYHSRGALQOIVRIKIEFHALE  
MSFOTRILSRVQLEEDTLAASEIKAAVYDAMRLTGOHAPVTEFIMSTFLLKPR  
VGSDFYLOENVRGEMAAKARPAVACAEVATGDEOTPCASASSKVLLELDFVAM  
YTYVHNIEVAAVVERSGRPDRDDEGLIEITOMPLISFPLVHRELARVALGGDVRPL  
RKPVCAADPAANIESPBGDALSEBIDQEGTOLSLQKRSKDKQVYIDRGSVERG  
KVGEDVYRTSSGVVRGQULTLFTVDEGVQAGSTIAPATPAENNSGNPLVEDDG  
TGSYSYPASPRSEHEABEDBEIIEFBEVDGANDDEBDALIELARKIGM"  
<6068..>6526  
/product="unknown"  
6068..6526  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAL78035.1"  
/db\_xref="GI:18656899"  
/translation="MVLONHQANRPPTPLSLTYSSRAVDYLSLVNMGADCFPLTGWPH  
MMNQPADQIGTSPGTSPRDP PFWHQQRDSQHAFOLAGTSLDYLVDMLGSEAEIP  
ARLYSSLSIHLRPSRSTMDSRVSSSLRAVSSSLRAVDSQKVAHQCLFPLP"

ORIGIN  
Query Match 1.3%; Score 53.8; DB 8; Length 6688;  
Best Local Similarity 54.9%; Pred. No. 0.024;  
Matches 106; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
Qy 2874 CTTGTGGCGCATCCCGACCCATTTCTCTAGTATATCAACGGTTGACGACTT 2933  
Db 4207 CGTAGTAGCCGTAACCCACGCGCTTCTCTGAGATGTCGGGTTGTAGCGAGTTGA 4266  
Qy 2934 CAAGCCGAGCTCCACAGCTCAAGAGTGAATCATCTTCTTAATAAGAGATGATACC 2993  
Db 4267 CGTGCCAGAGTTAGACATCTCAAGACGAGTGCTCTTCCAAAGCTGAGACAGGA 4326  
Qy 2994 GCTTGCTAAGAGCTATCTGTGAGAGACTAGACGCGGATATGAGCTGTGCTGGGA 3053  
Db 4327 CGTTCGAGCATGTGCTCCGATGAGTGCATGACGAGAGATGACTACTTGTACTGGGA 4386  
Qy 3054 TCCGAGATCGTC 3066  
Db 4387 CGAGCGACTCATC 4399

RESULT 7  
AX655393 2000 bp DNA linear PAT 22-MAR-2003  
LOCUS Sequence 5263 from Patent WO03000898.  
AX655393  
AX655393.1 GI:29158207

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
Plant genes involved in defense against pathogens



Qy	1991	CGACACCTCAAGTTGTTCTCCCGCGTATCCAGCTCCGCTGTGAGTAAGACATATGCCATTATGA	2050
Db	172	YGGACWCSCKSMNMWMSGCGCTYKRWKRSKYSWCCKKCYGSCCTKYSYTYRYCKMYK	113
Qy	2051	CATTGAGACCCCTCACACGATGACAGACACCAACGAGATCTTCTTACCTTCAGGCACTG	2110
Db	112	KSYYKCYCYCYWYWSYKRYMKKCMCSKSSGSMSCAYCSTSSISRMSMTYYAAKMGCG	53
Qy	2111	GCGAAGTGATGAATGACGGTGTAG	2135
Db	52	SSGMWRMSKSKMYSKYSCKYTG	28
RESULT 9			
LOCUS	A10243	1164 bp	DNA linear PAT 25-JAN-1994
DEFINITION	S.murinus DNA for glucose isomerase.		
LOCUS	A10243		
VERSION	A10243.1	GI:490202	
KEYWORDS	glucose isomerase.		
SOURCE	Streptomyces murinus		
ORGANISM	Streptomyces murinus		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.		
AUTHORS	Luiten,R.G.M., Quak,W.J., Schuurhuizen,P.W. and Mtabet,N.		
TITLE	Novel glucose isomerase enzymes and their use		
JOURNAL	Patent: EP 0351029-A 16 17-JAN-1990;		
FEATURES	GIST-BROCADES N.V.; PLANT GENETIC SYSTEMS, N.V		
source	location/Qualifiers		
	1..1164		
	/organism="Streptomyces murinus"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:33900"		
	1..>1164		
	/codon_start=1		
	/transl_table=1		
	/product="glucose isomerase"		
	/protein_id="CA00885.1"		
	/db_xref="GI:490203"		
	/db_xref="GOA:P37031"		
	/db_xref="UniProt/Swiss-Prot:P37031"		
	/translation="MSFQPTEDRFTFLMTKTVNGQRPFDATRPALDPVETQRLA		
	ELGAVYTFPHDDLIPGSSDTERESHKTRKRALDNGMVPATNLTPHVPKQK		
	GFTANDRVRRYALRKTIQNDILAELGATKYVAMGREGESGADVDRLDMRG		
	AFDLIGETVTAQGDILRPAIRPKNEPERGDLILPTVGALFIRLRPELPGVNPV		
	GHEWAGLNPFPHGIAQALMAKGLPHIDNGSGIKYQDILFFGADILRAFWLVDLIE		
	TAGEGSRHPIKPRTEPDFGVASAGCKRNYLILKDRAAAFRADPEVEALRARAR		
	LDQAQFPAAADGLALLADRAAFDFVDMAAAGMAFVHHDLQAMHLIGARG"		
ORIGIN			
Query Match	1.2%;	Score 51.6;	DB 6; Length 1164;
Best Local Similarity	47.8%;	Pred. No. 0.089;	
Matches 150;	Conservative 0;	Mismatches 164;	Indels 0; Gaps 0;
Qy	31	AGCCCCGTCGAGAAATCATTAACCCGGCTCAATAACGACTACAACCTGGGCTCCAGTGT	90
Db	16	ACCCCGGAGACAGTTTACCTTCGCTGTGGACCGTGGCTGGCAAGGAAGGACCCG	75
Qy	91	GTCGACAGACAACTCTGACCCCCACCGCCGAGAGAGACTGGCCGAGATGACGAGAT	150
Db	76	TTCGGCGACGCCACCCGCGCCGCTCGACCCGGTCAAGAGGTGACGCGCTGGCCGAG	135
Qy	151	TTCGTCGGCATGACAAAGTCTACAGAGACCTTGAACTTTCCTACTGCGGAGAGATGAC	210
Db	136	CTGGGCGCTTCAGAGTACCTTCACAGACGACGACTGATCCCTTGGGTCTTCGAC	195
Qy	211	TCCCTGAACACAGGCGAAGCAACTTCTTACATCGAGCGCAAGACTGCGAGTCTGAATGG	270
Db	196	ACGAGCGCGAGTCCGACATCAAGGCTTCCCGCCAGCGCCCTTGAGACCGACCGGATACG	255
Qy	271	GTGCCAAAGCCACGCGGACCTTGACACGCTTCCGTGATCCAAAGAACTTCCCGCGCGC	330

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	FEATURES
Db	256	GTGGCCATTGGCGACACCAACTTCTTACCCACCCCGCTTTCAAGAGACGGCGCTTACCC	315								
Qy	331	GCTACTGCGGCGCA 344									
Db	316	GCCACGACCGCGCA 329									
RESULT 10	AF443073	3524 bp DNA linear	PLN 03-DEC-2001								
LOCUS	Phomopsis sp. CMW 5588	putative RNA-dependent RNA polymerase									
DEFINITION	(rdp-1) gene, partial cds.										
ACCESSION	AF443073										
VERSION	AF443073.1	GI:17227101									
KEYWORDS	Phomopsis sp. CMW 5588										
SOURCE	Phomopsis sp. CMW 5588										
ORGANISM	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Diaporthales; Valsaceae; mitosporic Valsaceae; Phomopsis.										
REFERENCE	1 (bases 1 to 3524)										
AUTHORS	Preisig, O. and Wingfield, M. J.										
TITLE	Direct Submission										
JOURNAL	Submitted (04-NOV-2001) Forestry and Agricultural Biotechnology Institute, University of Pretoria, 72 Lunnon Street, Pretoria 0002, South Africa										
FEATURES	Location/Qualifiers										
SOURCE	1..3524										
	/organism="Phomopsis sp. CMW 5588"										
	/mol_type="genomic DNA"										
	/isolate="CMW 5588"										
	/db_xref="taxon:179291"										
	/country="South Africa"										
	<1..3524										
	/gene="rdp-1"										
	join(<1..2092,2148..3524)										
	/product="putative RNA-dependent RNA polymerase"										
	join(<1..2092,2148..3524)										
	/gene="rdp-1"										
	/note="RDP-1"										
	/codon_start=1										
	/product="putative RNA-dependent RNA polymerase"										
	/protein_id="AA138011.1"										
	/db_xref="GI:17227102"										
	/translation="HKTKHLPASLVKRVGSGVERGRVTVYFDINSFSLSSPSSVRYT YLAVQNVKTVQVNRFEVLVTPGPSVAMLLNLNSMEVLTLLSNTTTHLDPQVRLQLE TISGLLDEYSVNAQFLDKLNSPADARMMLEBMTLASIQLEPMKI PDQPKVLYHW PNAVPYATLVRAVAVTPTTIIYEKTCVELYTRKLYSDLNRPLRVQFTBEITGCG KIFESODSNKODNLIYRHWVMONGISIGDRHVRFLAFSSQPRENSQFECEDHTVY CDSIRMGSDGPRHTRISVSKFAARMGCGFTTRVONGISINIRKIIDIBSQTHGHTM PTDSVGKISNPFALINSERDLPBETPSCFQMRIGCGSVLYMVDVPSVSVHVRPSQE KFKVAVNVLITIKTFARSHATPNKQVLPVIALGVNSVSVSMDELDELREDEALAS MKABELLRQVDENQTTLTAEMVDITPMSKEPLMTLRLMKWILKRLKHKFAISV KKSMAFGVNDIEILGRHSQATEGRGNSIESLPQIFLVPEIGSGRSTTVNEVIT GICVAGNSLHPGDVAVADVDELRLHKNVVFPETGDRIDIPMSCGMDMGDDY FYVWDERLIPTEMHPPLNPDAGSSTLDKPADVTIENVTFFPAQYMNDSIGRLATAHA FAQADJQSGCVKPKCEILAKHSMAYDIKSGRPAMKXHLQRRKPHMREKSKSN YRQSLAIGKIYRIKIEFHAAYBMPDARILSRVQLSEADTLTAAKSIKATYDIAMRR LMGHEAIVTEFEIWTSTFILSRPVSQDYRLQENVELEALAKRFAECBEAVTGNV QTEFEPARELSVNLKEDRFVAAVMTYHNDVRAALRESRPNKVGSGSBEIOMPI SPEPLFRRLARVALGRGDVVRPLQKSWMLNRGSGOQTRBDVADAEFEIPDGLTAL PDALIEVETQQAQVGVNSKEGYTTERGSGADGELGDCVSTSGVVAHRCOMLAL PTGAEVEVQQAQSMPTDISPACQGLQASDVSNNSLVAVDSTGSSCPASPRSRRE VVEREEDIEIYBVDNAGYEEVEDALELARKLGI"										
ORIGIN	Query Match	1..28;	Score 51.4;	DB 8;	Length 3524;						
	Best Local Similarity	55.2%;	Pred. No. 0.1;								
	Matches 100;	Conservative 0;	Mismatches 81;	Indels 0;	Gaps 0;						

Db	1707	CGTGTGGCCGCAATCATCCATCCATCCGAGATGATCGAGTTGTGAGGACAGTA	1766
Qy	2934	CAAGCCAGAGCTCCACAGTCTCAAGAGATGTAATCATCTTCTCTAAGAGATGTA	2993
Db	1767	CGTCCCTGAGTTGTGACATCTGAAAGACGTGTCTGTCTTCCCAAGACTGGCCAGAGA	1826
Qy	2994	GCTTGTAAAGAGTATCTGTGTGAGACTACGACGGGATATGGCTGTGTCTGTGGA	3053
Db	1827	CATCCAGCATGTGTCTTGAGGTGACATGACGGGATGACTATTGTCTATTGGA	1886
Qy	3054	T 3054	
Db	1887	T 1887	
RESULT 11			
LOCUS	AF429315	125020 bp	DNA linear PRI 18-JAN-2002
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.		
ACCESSION	AF429315		
VERSION	AF429315.1	GI:17646244	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 125020)		
REFERENCE	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,		
REFERENCE	Ingberoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,		
REFERENCE	Potter,N.T., Ross,C.A. and Margolis,R.L.		
REFERENCE	A repeat expansion in the gene encoding junctophilin-3 is		
REFERENCE	associated with Huntington disease-like 2		
REFERENCE	Nat. Genet. 29 (4), 377-378 (2001)		
TITLE			
JOURNAL			
MEDLINE	21583737		
PUBMED	11694876		
REFERENCE	2 (bases 1 to 125020)		
REFERENCE	Holmes,S.E., Ingberoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.		
REFERENCE	Direct Submission		
REFERENCE	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical		
REFERENCE	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..125020		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="16"		
	/map="16q24.3: between D16S520 and WI-12410"		
	/note="Isolated from a patient with Huntington's		
	Disease-Like 2 (HDL2)"		
	complement (35581..35746)		
	/rpt_type=tandem		
	/rpt_unit="ctg"		
	complement (<36507..>36887)		
gene	/gene="JPH3"		
	/note="synonym: JP3"		
mRNA	complement (<36507..>36887)		
	/gene="JPH3"		
	/product="Junctophilin 3"		
CDS	complement (<36507..36887)		
	/gene="JPH3"		
	/note="component of the junctional complex between plasma		
	membrane and endoplasmic reticulum"		
	/codon_start=1		
	/product="Junctophilin 3"		
	/protein_id="AA140941.1"		
	/db_xref="GI:17646245"		
	/translation="MSSGGRFPNPDGSGYCGWEDGKAHGVCTGPKGGQGYTGSWS		
	HGPEVLGVTPSGNTYQGTWAGKRGHIGESKKNVYKGEWTHGKGRYGVRECSAG		
	NGAKYEGTWSNGLDGTYGTYSDG"		
ORIGIN			
Query Match	1.2%; Score 50.8; DB 9; Length 125020;		

Best Local Similarity 11.5%; Pred. No. 0.15;			
Matches 110; Conservative 407; Mismatches 428; Indels 9; Gaps 1;			
Qy	1698	GTGGCTCAGACAGGGGCAATCTCTTGTAGGCGCCGCAATGGCGGCTTTCTCGCA	1757
Db	51541	GAGCCSAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	51482
Qy	1758	AGATGCGGATACAGAAACCTCTCAGGAGTTCCAGCTCCGCGGAGAGCCGAAACC	1817
Db	51481	MMATACAMRWRAMMMSCMRKCTWSKMSHGRMRMYCWMARWYVYAGYSKSRWG	51422
Qy	1818	CATCATCAAGAGAGAGTCTTTTGGCCGAGACCGGCTTATGCTCCGACTGATGT	1877
Db	51421	GKGRSRSRSGMSWRTRGYSYSARCYMSKAGAGCKGAGYAKSAGRMWSYKMGAW	51362
Qy	1878	GTTCAAGACGAGATCTGTGCTCGGAGAGAACTGTAGACACGACGAGTCA	1937
Db	51361	YSCRMASKSRNTCTWSCCCTSCKCYMYMSWKYAGKCMYMRCTCMARMRAR	51302
Qy	1938	AGTTAGTCAATGCTGAGTGTGCTCTGCAATCGAACACACTTGGCAGCGACCT	1997
Db	51301	AAARBARMAAARAMMMMMWMMWRBHAAAMMAAAMMAAAMWYVYKCYMMSMMA	51242
Qy	1998	CAA-----GTTGTTCTCCGCTATCCAGCTCGCTGAGTAAACATATCCATTAT	2048
Db	51241	SWAMAMMMCTAAWYKKTTRMRMSKYSSSAYAVMYTMCYSSMYAARMGTYSG	51182
Qy	2049	GACATGAGACCTCACAGATGACAGACACACAGACGATCTTCTTCACTTACAGAC	2108
Db	51181	GKGRKMASTSCMCCRCCKMKKGRWSTTTTGMSSGKSRMTSMACMYKSRRK	51122
Qy	2109	TGCGAAGTATGATGACGCTGTAGCCGATGTCGGAAGCGTGGCCAAAGATACG	2168
Db	51121	TKSMATSGCMRWGAMSKGMRKYMSCRYKMGMRWMSGVKCSGVMSMGYSGRTSKY	51062
Qy	2169	CGATGTTCTGCTGGTGGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2228
Db	51061	CSATYGCSCSKAMKSKSTSKTSCTTSKYSLSMSRYSYSTCWMGRMKGMKSSCAS	51002
Qy	2229	GGAATGTGGTATGATGACCTTGACAGACAGGAGATGAGATGATGATGATGATGAT	2288
Db	51001	RSCKSYKMAAGSRRCGRGRMSCYWMSMKYSMAYSKARBRKCYKTSCKMYYSN	50942
Qy	2289	GTCCAGCGCAAGTGAATGCACTTCTGTGTAATCAACGTAACCTCGAAGTCCG	2348
Db	50941	HBKSHDSBGYSKYSRYBYVTGMMKDSVDSHMDMDYDKRMKYSSTGRYHBN	50882
Qy	2349	GACCGTGGCTTGAAGTGAAGTCACTGCTCAACCTACAGCTGTACTGCTGGA	2408
Db	50881	SAWMCMSMSRDSGYSGMVHMKSYWKSASKMYISGMKSSVYBMDCMSMYBSKSVR	50822
Qy	2409	AGATAGAGCCAGGAGCAAGGATGCAATGCGCCAGCAATCGGATCTTATCAACA	2468
Db	50821	MBRRMSGYRYRYKCAVMRYBHRSTRYSVYGGKSCWVRDVSRRSRNHTSMSSWRMC	50762
Qy	2469	TTTGAACAGCACTTACGACGCAAAAGCATGCTTGAATGCCAGTGAATTTGCCA	2528
Db	50761	RCRGARMAKSCAGKMSCVMTGMCVHSRCMSKMCSDCNRABDAHVCRCBYKBCA	50702
Qy	2529	ATGGGTTACAGAGATTAATTCAGTGGCGCACTCGAGTCAGCAGCGCGCTGTCTT	2588
Db	50701	DBCYKRDSVSCWBAHMYWVBHBSKSSRBGMHMHACHKSKSKSVKVMYVKKKCY	50642
Qy	2589	TCTTGTGGCTCTGACAGTCAAGAGAGACATGAACTTGTATGATGAACG	2642
Db	50641	WCRCSCWCSGWVHGRDRVYKRTARGGRRAMMYRRAKAYYDSSDMSWSG	50588
RESULT 12			
LOCUS	AB038265	1610 bp	DNA linear BCT 11-JUL-2001
DEFINITION	Streptomyces olivaceoviridis xyla gene for D-xylase isomerase,		
	complete cds.		

ACCESSION AB038265.1 GI:14669805  
VERSION D-xylose isomerase.  
KEYWORDS Streptomyces olivaceoviridis  
SOURCE Streptomyces olivaceoviridis  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE  
1 Kaneko, T.  
TITLE Streptomyces olivaceoviridis D-xylose isomerase (xy1A) gene  
JOURNAL Published Only in Database (2001)  
AUTHORS Kaneko, T.  
TITLE  
JOURNAL  
REFERENCE 2 (bases 1 to 1610)  
AUTHORS Kaneko, T.  
TITLE Direct Submission  
JOURNAL Submitted (10-FEB-2000) Takahiro Kaneko, Akita Research Institute of Food and Brewing, Bioengineering Div., 4-26 Sanuki, Araya-nachi, Akita, Akita 010-1623, Japan  
(E-mail:kaneko@ake.aitf.pref.akita.jp, Tel:+81-18-888-2001)  
Location/Qualifiers

FEATURES  
source  
1..1610  
/organism="Streptomyces olivaceoviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1921"  
268..1434  
/gene="xy1A"  
268..1434  
/gene="xy1A"  
/codon\_start=1  
/transl\_table=11  
/product="D-xylose isomerase"  
/protein\_id="BAB62014.1"  
/db\_xref="GI:14669806"  
/translation="MSYQPTPEDRFTFGILMTVGQSDPFGDTRRLDPEVYQRLA  
ELGAGVTFHDDLIIPGSSDTERESHIKFRQALDAGTMVPAATNLTLPVFKG  
AFANDVDVRYALRKTIRINDLAELGATVYAMWGBEAGSAAVDVSALDRME  
AFDLGVTYSQGVLFPALEPKNEPRGDIILPTVGHALAFIERLPELYGVNPEV  
GHEOMAGLNPFGHIAQALWAGKLFHIDNGSGIKYPODLRFAGDRLSALMVDLLE  
SAGEGPRHDFRPTREDLDGVMAAAGCMRYLILKERAAFRADPVOALRASR  
LDQLAQPTAADGLEDLADRAAFEDPVEAAAGMFAERLDQALMHLGARG"

ORIGIN  
Query Match 1.2%; Score 50; DB 1; Length 1610;  
Best Local Similarity 47.5%; Pred. No. 0.23;  
Matches 149; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 31 AGCCCGTCGAGAAATCATTAACCGGCTCAATTAACGACTCAACCTGGGCTTCAAGTGT 90  
DB 283 ACCCCGAGGACAGGTTCACTGCGACTGCGACCTGCGGCAAGGAGGACCCG 342  
QY 91 GTGCGACAGCAACTTTCACACCCCGCGGGAAGAGCTGGCCGAGAGTGAAGAGAT 150  
DB 343 TTCGGCGACCGCACCGCGCGCCCTGACCCGGTGCAGACGGTGCAGCGCTGCGGAA 402  
QY 151 TTCGGTGCATGACAAGATCTACAGAGCCCTGAACCTTCTCTACTGCGGAGAGATGAC 210  
DB 403 CTCGGTGCACGAGAGTACCTTTCACAGACGACGACCTGATCCCTTCGTTCTGCGAC 462  
QY 211 TCCTGAAACAGGAGAAAGCACTTCTTCATCGAGCCAAAGCTGCGAGTCAAGTGG 270  
DB 463 ACCGAGCGGAGTGCACATCAAGCGGTTCCGCGAGGCGCTGACGCGACCGGATGACC 522  
QY 271 GTGCGCAAAGCCAGCGCGACCTGACACGCTTCGTTGTCAGAGAACTCCCGCGGCC 330  
DB 523 GTTCGAGTGCACCAAGCACTTTCACGACCCCGTTTCAGAGACGGCGCTTACG 582  
QY 331 GCTACTGCGGCGCA 344  
DB 583 GCCAACGACCGCA 596

RESULT 13  
AF170068 1952 bp DNA linear BCT 12-AUG-1999  
LOCUS

DEFINITION Streptomyces chibaensis D-xylose isomerase (xy1A) gene, complete cds.  
ACCESSION AF170068  
VERSION AF170068.1 GI:5731284  
KEYWORDS Streptomyces chibaensis  
SOURCE Streptomyces chibaensis  
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE  
1 (bases 1 to 1952)  
AUTHORS Joo, G.-J., Shin, J.H., Heo, G.-Y., Park, H.D. and Rhee, I.K.  
TITLE Streptomyces chibaensis J-59 xy1A  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1952)  
AUTHORS Joo, G.-J., Shin, J.H., Heo, G.-Y., Park, H.D. and Rhee, I.K.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUL-1999) Department of Agricultural Chemistry, College of Agriculture, Kyungpook National University, Buk-gu, Sanjuk-dong, 1372, Taegu 702-701, Korea  
Location/Qualifiers

FEATURES  
source  
1..1952  
/organism="Streptomyces chibaensis"  
/mol\_type="genomic DNA"  
/strain="J-59"  
/db\_xref="taxon:67286"  
552..1718  
/gene="xy1A"  
552..1718  
/gene="xy1A"  
/codon\_start=1  
/transl\_table=11  
/product="D-xylose isomerase"  
/protein\_id="AAD48850.1"  
/db\_xref="GI:5731285"  
/translation="MSYQPTPEDRFTFGILMTVGQSDPFGDTRRLDPEVYQRLA  
ELGAGVTFHDDLIIPGSSDTERESHIKFRQALDAGTMVPAATNLTLPVFKG  
AFANDVDVRYALRKTIRINDLAELGATVYAMWGBEAGSAAVDVSALDRME  
AFDLGVTYSQGVLFPALEPKNEPRGDIILPTVGHALAFIERLPELYGVNPEV  
GHEOMAGLNPFGHIAQALWAGKLFHIDNGSGIKYPODLRFAGDRLSALMVDLLE  
SAGEGPRHDFRPTREDLDGVMAAAGCMRYLILKERAAFRADPVOALRASR  
LDQLAQPTAADGLEDLADRAAFEDPVEAAAGMFAERLDQALMHLGARG"

ORIGIN  
Query Match 1.2%; Score 50; DB 1; Length 1952;  
Best Local Similarity 47.5%; Pred. No. 0.23;  
Matches 149; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 31 AGCCCGTCGAGAAATCATTAACCGGCTCAATTAACGACTCAACCTGGGCTTCAAGTGT 90  
DB 567 ACCCCGAGGACAGGTTCACTGCGACTGCGACCTGCGGCAAGGAGGACCCG 626  
QY 91 GTGCGACAGCAACTTTCACACCCCGCGGGAAGAGCTGGCCGAGAGTGAAGAGAT 150  
DB 627 TTCGGCGACCGCACCGCGCGCCCTGACCCGGTGCAGACGGTGCAGCGCTGCGGAA 686  
QY 151 TTCGGTGCATGACAAGATCTACAGAGCCCTGAACCTTCTCTACTGCGGAGAGATGAC 210  
DB 687 CTCGGTGCACGAGAGTACCTTTCACAGACGACCTGATCCCTTCGTTCTGCGAC 746  
QY 211 TCCTGAAACAGGAGAAAGCACTTCTTCATCGAGCCAAAGCTGCGAGTCAAGTGG 270  
DB 747 ACCGAGCGGAGTGCACATCAAGCGGTTCCGCGAGGCGCTGACGCGCATGACC 806  
QY 271 GTGCGCAAAGCCAGCGCGACCTGACACGCTTCGTTGTCAGAGAACTCCCGCGGCC 330  
DB 807 GTCCGATGCGCACCAAGCACTTTCACGACCCCGTTTCAGAGACGGCGCTTACG 866  
QY 331 GCTACTGCGGCGCA 344  
DB 867 GCCAACGACCGCA 880

RESULT 14

AR145904  
LOCUS AR145904 3731 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 1 from patent US 6218142.  
ACCESSION AR145904  
VERSION AR145904.1 GI:15109093  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 3731)  
TITLE Mueseneger M., Riedel L., Schiebel W. and Sanger H. L.  
Nucleic acid molecules encoding polypeptides having the enzymatic  
activity of an RNA-directed RNA polymerase (RDRP)  
Patent: US 6218142-A 1 17-Apr-2001;  
FEATURES  
source Location/Qualifiers  
1..3731  
/organism="Unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 1.2%; Score 49; DB 6; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.43;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGTATGCTCTCGTGGCGCCGATCCCGACCCATTCCCTAAGTGA 2909  
DB 2425 TAATTTCAATTCGAAGGAAATGCTGTTGCAAAAATCCATGCTTGATCCTGGTGA 2484  
QY 2910 TATCCAAGGGTTGAGCAGCTTCAAGCAGAGCTCCACAGTCTCAAGATGTAATCAT 2969  
DB 2485 TATTCGCTTTTAAAGCTGTAATGTTGAGCGCTCACCACATGTAATTTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATCGCTTGTCTAAGACCTATCTGTGAGACTAAGACGG 3029  
DB 2545 ATTCCTCAGAAAAGAAAAGACCTCATCCGAATGATGTTGCGAGTGTATTTGATGG 2604  
QY 3030 CGATATGGCTGGTGTCTGCTGGATCCGAGATGTC 3066  
DB 2605 GGATATCTACTTTGTTGCTGGATCAAGACATGATC 2641  
RESULT 15  
LERDRP 3731 bp mRNA linear PLN 18-DEC-1998  
LOCUS LERDRP  
DEFINITION L.esculentum mRNA for RNA-directed RNA polymerase.  
ACCESSION Y10403  
VERSION Y10403.1 GI:4038591  
KEYWORDS RDRP gene; RNA-directed RNA polymerase.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum; Streptophyta; Embryophyta; Tracheophyta;  
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE  
AUTHORS 1  
TITLE Schiebel W., Pelissier T., Riedel L., Thalmair S., Schiebel R.,  
Kemper D., Lottspeich F., Sanger H. L. and Mueseneger M.  
Isolation of an RNA-directed RNA polymerase-specific cDNA clone  
from tomato  
JOURNAL Plant Cell 10 (12), 2087-2101 (1998)  
MEDLINE 99055198  
PUBMED 9836747  
REFERENCE 2 (bases 1 to 3731)  
AUTHORS Mueseneger M.  
TITLE Direct Submission  
JOURNAL Submitted (08-JAN-1997) M. Mueseneger, Max-Planck-inst. fuer  
Biochemie, Viroidforschung, Am Klopferspitz 18a,  
Planegg-Martinsried, 82152, FRG  
FEATURES  
source Location/Qualifiers  
1..3731  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/strain="Rutgers"  
/db\_xref="taxon:4081"

gene 1..3731  
/gene="RDRP"  
CDS 194..3538  
/gene="RDRP"  
/EC\_number="2.7.7.48"  
/codon\_start=1  
/product="RNA-directed RNA polymerase"  
/protein\_id="CAA71421.1"  
/db\_xref="GI:4038592"  
/db\_xref="GOA:Q9ZR58"  
/db\_xref="UniProt/TREMBL:Q9ZR58"  
/translation="MKTITQVRFYLLSAYVSKFLKTYGTGYCALFVSKSGS  
RAFAKQFADNISADKXITLANNRLYFGSSYLKAMWKKTDIVQLRAYVQMDITLNF  
GCQISDDKPAVLGSTEVSIOFGIGLKKFPFLLSGSADYKLQSYENIQOVILHRYG  
ONAFLLIOLFGAPRIYKRLNENSCYFPKPEPDQVARTDDPSPWIGSSLCLEPR  
RGVRLPNFESSFPHVAERENNTLOTGTFPFPQKALVNPVQPGGISIPYILKPI  
SLVVGCTPGPALVYFPRILVDRRRNACIEHLAEKLYVKECCYDPRMTTEQD  
GYLKGQPKPSITLDDGLVYRRLVTPCKVYFCGPEVNSNRYLVKYSIEDINFL  
RVSEVDEWEKLYSTDLLPKASTGSGVRTNIYERILSTLRKGFVIGDKRFEPLAFSS  
QLRDNVMMFASRPGLJLANDIRAMQDFSOIKNVAKYAARLQSPGSSRETLVLRHE  
IEVTPVKVHTGTSYVPSDGIKISGDFARVASKCGLOTPSAFOIRYGVKGVAVD  
PDSMKLSLRKMSKYSNDIKLDVGMKGYOCPYLNROLITLSTLGYKDEVLGKQ  
KEAVDQDLIDLHDSLKAQBALHMSPEGENTILKAMLCGYVDPDAEPLSMITQFRA  
SKLDDRTRSRIFIPNGRTMGLDSESRFLIEGVQVYQTTGACGEFSDDLHFNRS  
STGNSFILGNVAVAKNPLHREDIRVLAVNRAVLAHHVDVCFVQKRRPHPNCS  
GISLDDGDIYFCVMDQDMPPRQVQPMVEYPPASIQLDHVDVTLIEVEYPTNYVNDL  
GIANAHVVFADREPDMAISDCKLAELFSLAVFPKPAELPSQURPKXPYDFM  
DKPDKTSYISERVIGLFRKVKDKAPQASSIATFTDVARRSYDADAEVDFEDYIDE  
APFKTEYNKGNLMDYVIGKTEARILSGGIMKSKTDFRRKDAEISVAARALKE  
APAFRRNDIDIDMLPKASAMHYVTHPYPMGCIYNGLGRAHFIISPMCVQDLQIK  
KDKARRPVLNLSLRQLSHRLVLEK"  
ORIGIN  
Query Match 1.2%; Score 49; DB 8; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.43;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGTATGCTCTGCGGCGATCCCGACCCATTCCCTAAGTGA 2909  
DB 2425 TAATTTCAATTCGAAGGAAATGCTGTTGCAAAAATCCATGCTTGATCCTGGTGA 2484  
QY 2910 TATCCAAGGGTTGAGCAGCTTCAAGCAGAGCTCCACAGTCTCAAGATGTAATCAT 2969  
DB 2485 TATTCGCTTTTAAAGCTGTAATGTTGAGCGCTCACCACATGTAATTTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATCGCTTGTCTAAGACCTATCTGTGAGACTAAGACGG 3029  
DB 2545 ATTCCTCAGAAAAGAAAAGACCTCATCCGAATGATGTTCTGGAGATGATTTGATGG 2604  
QY 3030 CGATATGGCTGGTGTCTGCTGGATCCGAGATGTC 3066  
DB 2605 GGATATCTACTTTGTTGCTGGATCAAGACATGATC 2641  
Search completed: August 11, 2005, 00:49:20  
Job time : 17526 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 12:28:07 ; Search time 2057 Seconds

(without alignments)  
12104.245 Million cell updates/sec

Title: US-09-913-878a-1\_COPY\_2447\_6652

Perfect score: 4206  
Sequence: 1 atgaacctctactctctag.....tcacaggggaatggcattat 4206

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001s:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4206	100.0	8045	3	AAA65171
2	4202.8	99.9	4206	12	ADN40832
3	3078	73.2	3078	12	ADN40834
4	93.2	2.2	615	8	ABZ54843
5	52.4	1.2	2000	8	ADA71938
6	51.8	1.2	2000	8	ADA71938
7	51.6	1.2	1164	4	AAQ03159
8	49	1.2	3731	4	AAI17845
9	49	1.2	3731	4	AAI17837
10	49	1.2	3731	4	AAI17837
11	45.6	1.1	495	9	ACH16195
12	45.6	1.1	2710	4	AAI17837
13	44	1.0	282	6	ABL71773
14	44	1.0	1050	4	AAI17837
15	44	1.0	1318	5	AAI17837
16	44	1.0	1540	3	AAI17837
17	44	1.0	1555	4	AAI17837
18	44	1.0	1831	6	AAI17837
19	44	1.0	1926	5	AAI17837
20	44	1.0	1959	5	AAI17837

21	44	1.0	3170	4	AAK85216
22	44	1.0	3364	4	AAK85217
23	44	1.0	3901	3	AAK63740
24	44	1.0	94310	11	ACN44662
25	42.4	1.0	506	12	ACH87424
26	42.4	1.0	1168	10	ADC87372
27	42.2	1.0	594	5	AAI17837
28	42	1.0	1167	1	AAI17837
29	41.2	1.0	1485	4	AAI17837
30	41.2	1.0	1485	4	AAI17837
31	41.2	1.0	1535	4	AAI17837
32	41.2	1.0	1535	8	AAI17837
33	41.2	1.0	1535	8	AAI17837
34	41.2	1.0	1535	8	AAI17837
35	41.2	1.0	1535	8	AAI17837
36	41.2	1.0	1535	8	AAI17837
37	41.2	1.0	1535	8	AAI17837
38	41.2	1.0	1535	8	AAI17837
39	41.2	1.0	1535	8	AAI17837
40	41.2	1.0	1535	8	AAI17837
41	41.2	1.0	1535	8	AAI17837
42	41.2	1.0	1535	8	AAI17837
43	41.2	1.0	1535	8	AAI17837
44	41.2	1.0	1535	8	AAI17837
45	41.2	1.0	1535	8	AAI17837

## ALIGNMENTS

RESULT 1	AAA65171	standard; DNA; 8045 BP.
XX	AAA65171;	
AC	16-NOV-2000	(first entry)
DT	16-NOV-2000	(first entry)
XX	Neurospora crassa qde-1 gene.	
DE	Neurospora crassa.	
OS	Neurospora crassa.	
XX	Key	Location/Qualifiers
FT	CDS	2447..6655
FT		/*tag= a
FT		/product= "QDE-1"
XX	MO200050581-A2.	
PD	31-AUG-2000.	
XX	16-FEB-2000; 2000MO-IT000048.	
PR	22-FEB-1999; 99IT-RM000117.	
XX	(UTRO-) UNIV ROMA LA SAPIENZA.	
XX	Macino G, Cogoni C;	
XX	WPI; 2000-579171/54.	
DR	P-85DB; AAB13956.	
XX	Novel polynucleotide encoding a polypeptide which has a silencing	
PT	activity and comprising a RNA-dependent RNA polymerase domain.	
XX	Claim 1; Page 31-43; 48pp; English.	
XX	The present sequence is the Neurospora crassa qde-1 gene. This gene has	
CC	silencing activity. The qde-1 gene was isolated by mutational analysis of	
CC	an al-1 transgenic strain. This strain had an albino phenotype resulting	
CC	from post-transcriptional silencing of the endogenous al-1 gene.	

CC Reversion of this phenotype indicated a mutation in a silencing gene. The  
CC silencing gene, qde-1, could then be isolated. Modulation of qde-1  
CC expression may be used to inactivate genes and to silence suppression of  
CC genes

XX Sequence 8045 BP; 2089 A; 2180 C; 1950 G; 1826 T; 0 U; 0 Other;

Query Match 100.0%; Score 4206; DB 3; Length 8045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAACCTTATTACTCTTAGAGAGAGATAGCCCGCTGAGAGAAATCATTAACCGGCTC 60
DB 2447 ATGAACCTTATTACTCTTAGAGAGAGATAGCCCGCTGAGAGAAATCATTAACCGGCTC 2506
QY 61 AATTAACGACTACAACTCTGGGCTTCCAGTGTGTGCGACACAACTTACCCGCCACCGC 120
DB 2507 AATTAACGACTACAACTCTGGGCTTCCAGTGTGTGCGACACAACTTACCCGCCACCGC 2566
QY 121 CGGAGAGAGCTGGCCGAGAGTACAGAGATTTGGGTGCCATGACAAAGATCTACAGACC 180
DB 2567 CGGAGAGAGCTGGCCGAGAGTACAGAGATTTGGGTGCCATGACAAAGATCTACAGACC 2626
QY 181 CTGAACCTTCTACTAGCGGAGAGATGACTCCCTGAACGAGGAGAGGCACTTCTTC 240
DB 2627 CTGAACCTTCTACTAGCGGAGAGATGACTCCCTGAACGAGGAGAGGCACTTCTTC 2686
QY 241 ATGAGGCGCAAAGCTGCGAGCTCGAATCGGATGCCAAAGCCGCGGACCCCTGACAGC 300
DB 2687 ATGAGGCGCAAAGCTGCGAGCTCGAATCGGATGCCAAAGCCGCGGACCCCTGACAGC 2746
QY 301 CTTCCTGCTGTCTCAAGAACTCCCGCGCGCTACTGCGCGGCAACAATGGGCAATTGCG 360
DB 2747 CTTCCTGCTGTCTCAAGAACTCCCGCGCGCTACTGCGCGGCAACAATGGGCAATTGCG 2806
QY 361 ACTGTGTGCTCGAGGCTTAATAGGTTATGCACTCCCAATACAGACGAGTGC 420
DB 2807 ACTGTGTGCTCGAGGCTTAATAGGTTATGCACTCCCAATACAGACGAGTGC 2866
QY 421 ACGTTTGGCAGAACTCTAAGCGGCGCAAGTGCGCTGAGCGCGCAACTCTTACCAACCC 480
DB 2867 ACGTTTGGCAGAACTCTAAGCGGCGCAAGTGCGCTGAGCGCGCAACTCTTACCAACCC 2926
QY 481 AAACGCGAAGATGAGCCCGCAATGTCACTTTGCTGATTCGCCCAAGCGCTGGTGAAT 540
DB 2927 AAACGCGAAGATGAGCCCGCAATGTCACTTTGCTGATTCGCCCAAGCGCTGGTGAAT 2986
QY 541 CGGCTGCGCAGAGTCTCTCATTTACGCGCGCGCGATACCCCTAAAGTCCCGGATCCA 600
DB 2987 CGGCTGCGCAGAGTCTCTCATTTACGCGCGCGCGATACCCCTAAAGTCCCGGATCCA 3046
QY 601 GTGAATACCGGTTCCAAACGACATCTCTGAGAGTGAAGATCTCAATCAGTGACCAAG 660
DB 3047 GTGAATACCGGTTCCAAACGACATCTCTGAGAGTGAAGATCTCAATCAGTGACCAAG 3106
QY 661 CGGCGCCAAAGGCAAGCTGTCTGATATGTTGCGGCTGCGCGCGCGCGCTGGTCAAT 720
DB 3107 CGGCGCCAAAGGCAAGCTGTCTGATATGTTGCGGCTGCGCGCGCGCGCTGGTCAAT 3166
QY 721 GCGAGCGCTTTGACAAAGGTAACGACTCGAAGAGATGCCAATACGAGAGATCCCAAGGCG 780
DB 3167 GCGAGCGCTTTGACAAAGGTAACGACTCGAAGAGATGCCAATACGAGAGATCCCAAGGCG 3226
QY 781 ACAGGTATAGACGAGCGGACAGGTGATTTCTTTGATATCATCTCAAGGCACTTCTAT 840
DB 3227 ACAGGTATAGACGAGCGGACAGGTGATTTCTTTGATATCATCTCAAGGCACTTCTAT 3286
QY 841 GGTTCAGTGTCTTACAGCGCTTGGCGTCAACATTCAGACATCCAGAGTAGTTTGG 900
DB 3287 GGTTCAGTGTCTTACAGCGCTTGGCGTCAACATTCAGACATCCAGAGTAGTTTGG 3346
QY 901 GGTTCCTCTTCAAGCCGAGAGAGAGAGCGGCTGTGATGCGCAGGTCTTTGAGGCTGGA 960
```

```
DB 3347 GGTCTCTCTTCAAGCCGAGAGAGAGAGCGGCTGTGATGCCACGGTCTTTAGGCTGGA 3406
QY 961 CACTGTATGAGTCTCTTAGCAAAAGAGAACAAACAGTCCGACATAGATAACGAGCC 1020
DB 3407 CACTGTATGAGTCTCTTAGCAAAAGAGAACAAACAGTCCGACATAGATAACGAGCC 3466
QY 1021 CTTTCAATGCTTCCCAAGGTTGAACCTTGTGAGCATTTACTATGATGCTTTCCAGT 1080
DB 3467 CTTTCAATGCTTCCCAAGGTTGAACCTTGTGAGCATTTACTATGATGCTTTCCAGT 3526
QY 1081 TCCGCGCGGAGGCGGCAATTTCCGAGCGGAGTCCGCTCAAAATGAGTCTGGCGAGAA 1140
DB 3527 TCCGCGCGGAGGCGGCAATTTCCGAGCGGAGTCCGCTCAAAATGAGTCTGGCGAGAA 3586
QY 1141 GAAAGCGCTGATCTCAGGTTCAAGTTCAATGCCCGGTGGTTGCACTCGGCTGAGAAAT 1200
DB 3587 GAAAGCGCTGATCTCAGGTTCAAGTTCAATGCCCGGTGGTTGCACTCGGCTGAGAAAT 3646
QY 1201 ATTTGGCCGAATTTTCCCAATGAGCTACACGAAGTCTCTCGCTGTTGATGGAAATT 1260
DB 3647 ATTTGGCCGAATTTTCCCAATGAGCTACACGAAGTCTCTCGCTGTTGATGGAAATT 3706
QY 1261 ACCAGACTTTTATGACTGCAAAAGTGAATTGGAAGAGAGAGGCTGGGCTTAAAGTAC 1320
DB 3707 ACCAGACTTTTATGACTGCAAAAGTGAATTGGAAGAGAGAGGCTGGGCTTAAAGTAC 3766
QY 1321 GACCTTCTGTGTCTACCGGCGCGATGTCAAGATATGTGAAGAATCTCTTACCGGCTT 1380
DB 3767 GACCTTCTGTGTCTACCGGCGCGATGTCAAGATATGTGAAGAATCTCTTACCGGCTT 3826
QY 1381 GATGCTTTCCGTGTGAACCTTTTCCAGAAAGCGGCCCAAGAGTGTTCGTGACCGCA 1440
DB 3827 GATGCTTTCCGTGTGAACCTTTTCCAGAAAGCGGCCCAAGAGTGTTCGTGACCGCA 3886
QY 1441 ATGAGCGGCAACTTTGAGAGCAAAAGGTAGCGGCTGTTCTCTGCTGTTCTAGACTAC 1500
DB 3887 ATGAGCGGCAACTTTGAGAGCAAAAGGTAGCGGCTGTTCTCTGCTGTTCTAGACTAC 3946
QY 1501 AATCCGAGCAACTCGGCTTACTGCGCCCTTTACTTTGTAAGCTGAAGCGCTCATGTTTC 1560
DB 3947 AATCCGAGCAACTCGGCTTACTGCGCCCTTTACTTTGTAAGCTGAAGCGCTCATGTTTC 4006
QY 1561 GAGCAGGCTGTGTGATCACCGGTGGTTCGTGATAGTGTTCGAGATCCTTAAT 1620
DB 4007 GAGCAGGCTGTGTGATCACCGGTGGTTCGTGATAGTGTTCGAGATCCTTAAT 4066
QY 1621 CGGTGCGCTACGAGCAACGAGCCCAAGTGAACGCGGTGTGAGCAAAACAACGAGTGG 1680
DB 4067 CGGTGCGCTACGAGCAACGAGCCCAAGTGAACGCGGTGTGAGCAAAACAACGAGTGG 4126
QY 1681 GTGGAAGATCATCCAGTGGCTCACGATGGGCAACATTTCTTTGTAGGCGCCCAATGG 1740
DB 4127 GTGGAAGATCATCCAGTGGCTCACGATGGGCAACATTTCTTTGTAGGCGCCCAATGG 4186
QY 1741 CGCGCTTTCTTGGCCAAAGATGCCGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 1800
DB 4187 CGCGCTTTCTTGGCCAAAGATGCCGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 4246
QY 1801 GCGGAGACCCGAAACCCCATCATCAAGGAGAGTCACTTCTTTCCGAGACCGGCAATT 1860
DB 4247 GCGGAGACCCGAAACCCCATCATCAAGGAGAGTCACTTCTTTCCGAGACCGGCAATT 4306
QY 1861 ACGTTCGACATGATGTGTTCAAGAGAGATCTGTGTTCCGCGAGAGAACTCTGAGAG 1920
DB 4307 ACGTTCGACATGATGTGTTCAAGAGAGATCTGTGTTCCGCGAGAGAACTCTGAGAG 4366
QY 1921 CAACGAGCCGAGTTCAAGATGATCAATGCTGAGACTGCTCTGCAACCAAC 1980
DB 4367 CAACGAGCCGAGTTCAAGATGATCAATGCTGAGACTGCTCTGCAACCAAC 4426
QY 1981 ACTTGGCAGCCGCACTCAAGTGTCTCCCGATACCACTGGTGTGATGAAGATAT 2040
DB 4427 ACTTGGCAGCCGCACTCAAGTGTCTCCCGATACCACTGGTGTGATGAAGATAT 4486
```



2041 GCCATTATGACATTGAGCCCTCACAGATGAGACCAAGACCGATCTTTCACT 2100  
4487 GCCATTATGACATTGAGCCCTCACAGATGAGACCAAGACCGATCTTTCACT 4546  
2101 TCAAGCAGCTGGCGAAGTGAATGACGGTGTAGCCGCATGTGCGAAGCGTGGCCAG 2160  
4547 TCAGGCACTGGCGAAGTGAATGACGGTGTAGCCGCATGTGCGAAGCGTGGCCAG 4606  
2161 AGGATAGCCGATGTTCTCGGTTTGGGTGATGTCCTCTGTGTGCAAGGGCGTTGGT 2220  
4607 AGGATAGCCGATGTTCTCGGTTTGGGTGATGTCCTCTGTGTGCAAGGGCGTTGGT 4666  
2221 TCGGCGAAGGGAATGTGGGTTATCGACGTTGACGACAGCGATGAGATTGGATGAG 2280  
4667 TCGGCGAAGGGAATGTGGGTTATCGACGTTGACGACAGCGATGAGATTGGATGAG 4726  
2281 ACATACCCGCTCCAGCGCAAGTGGAAATGCGACTTGTGATTAACATCAACGTACCTC 2340  
4727 ACATACCCGCTCCAGCGCAAGTGGAAATGCGACTTGTGATTAACATCAACGTACCTC 4786  
2341 GAAGTCGGAGCGTGGCTTCTGAATGAACTGACCTGCTCAACCTACAGCTTACCT 2400  
4787 GAAGTCGGAGCGTGGCTTCTGAATGAACTGACCTGCTCAACCTACAGCTTACCT 4846  
2401 GTCCTGGAAGTAAAGCAGGGAAGAGTGAAGTGGCCAGGCAATCGGTGACCTT 2460  
4847 GTCCTGGAAGTAAAGCAGGGAAGAGTGAAGTGGCCAGGCAATCGGTGACCTT 4906  
2461 ATCAACGATTTTGAACGACAGTTCAGGAGCAAAAGCATGCTTTGAATCCGCGAGTGA 4966  
4907 ATCAACGATTTTGAACGACAGTTCAGGAGCAAAAGCATGCTTTGAATCCGCGAGTGA 4966  
2521 TTTTGGCCATGAGGTTTACGAGATTTATTCAGTGGCCGCACTCGAGTACGACGCGCT 2580  
4967 TTTTGGCCATGAGGTTTACGAGATTTATTCAGTGGCCGCACTCGAGTACGACGCGCT 5026  
2581 GTGCTTTTCTTGTGGGCTACCTGACGATGAGGAGACATCGAATCTTTGATGAC 2640  
5027 GTGCTTTTCTTGTGGGCTACCTGACGATGAGGAGACATCGAATCTTTGATGAC 5086  
2641 AGTGGGTTGATCCCAAGAACAAAGTACTTGCAGACATCGCTGGGATCTTCAAAAG 2700  
5087 AGTGGGTTGATCCCAAGAACAAAGTACTTGCAGACATCGCTGGGATCTTCAAAAG 5146  
2701 CGGAAATGTGACAGTGTGAATCCAGCTGAACATCCGTGCGTCCGATCAGCATACAT 2760  
5147 CGGAAATGTGACAGTGTGAATCCAGCTGAACATCCGTGCGTCCGATCAGCATACAT 5206  
2761 TACATGATTTGCGGATTTCTGGGGTGTGCTTGAAGAAAATGAGTTCAATGCGGATTC 2820  
5207 TACATGATTTGCGGATTTCTGGGGTGTGCTTGAAGAAAATGAGTTCAATGCGGATTC 5266  
2821 TCAAAATTCAGAGACGAGAGAGATCTTTTACATCTCTATCGACTGTATGTCCTG 2880  
5267 TCAAAATTCAGAGACGAGAGAGATCTTTTACATCTCTATCGACTGTATGTCCTG 5326  
2881 GCGGATTCGCCAGCCCATTTTCTAGTATATCCAAACGGGTTGAGAGCTTTCAAGCA 2940  
5327 GCGGATTCGCCAGCCCATTTTCTAGTATATCCAAACGGGTTGAGAGCTTTCAAGCA 5386  
2941 GAGCTCCAGCTCAAGGATGTAATCATCTTCTACTATAAGGAGATGTAACCGCTTGT 3000  
5387 GAGCTCCAGCTCAAGGATGTAATCATCTTCTACTATAAGGAGATGTAACCGCTTGT 5446  
3001 AAGAGCTATCTGTGAGACTACGACGCGATATGCGCTGCTGCTGGATCCGAG 3060  
5447 AAGAGCTATCTGTGAGACTACGACGCGATATGCGCTGCTGCTGGATCCGAG 5506  
3061 ATCGTCATGCTTGTGTAATGCGGAAATGCTCTGAGCCGACCTGTCTAGGTAACCTA 3120  
5507 ATCGTCATGCTTGTGTAATGCGGAAATGCTCTGAGCCGACCTGTCTAGGTAACCTA 5566

3121 AAGAGGCAAAACGACTTTCAAACTTATGCGCTCACAGGCAAGGCTCAGCGGCC 3180  
5567 AAGAGGCAAAACGACTTTCAAACTTATGCGCTCACAGGCAAGGCTCAGCGGCC 5626  
3181 AAGAGGCACTACATACGATATGATCCAGAGAGCTTCCATTTCCGCTGCAAGCCAC 3240  
5627 AAGAGGCACTACATACGATATGATCCAGAGAGCTTCCATTTCCGCTGCAAGCCAC 5686  
3241 TTTCTGGGCAATGACCTAACCTAACAAAGGCTGTTACATCAACATAGTGTCT 3300  
5687 TTTCTGGGCAATGACCTAACCTAACAAAGGCTGTTACATCAACATAGTGTCT 5746  
3301 AAGAGCCGCGCATCATTTCTTATGTTCACTGCTGGGAAACCTCGTGCATCAGACAA 3360  
5747 AAGAGCCGCGCATCATTTCTTATGTTCACTGCTGGGAAACCTCGTGCATCAGACAA 5806  
3361 GGTATGTCCTTAAACGAGACAGCTGGGCTCAATTTGCTGAGGAACTGCTTGGCGTGA 3420  
5807 GGTATGTCCTTAAACGAGACAGCTGGGCTCAATTTGCTGAGGAACTGCTTGGCGTGA 5866  
3421 TTTGCTTCTGACCCCAATGTAACAAGGAGCAGTTGGCTCGGGGCGGAGAGCTTAC 3480  
5867 TTTGCTTCTGACCCCAATGTAACAAGGAGCAGTTGGCTCGGGGCGGAGAGCTTAC 5926  
3481 CACATTATGACTTACCTGAATTTCTCCATGCGCAGGCTCGATTTGACAAAGAACTGAA 3540  
5927 CACATTATGACTTACCTGAATTTCTCCATGCGCAGGCTCGATTTGACAAAGAACTGAA 5986  
3541 GCTTCCCAATGCGCATGAAAGCGGCCAGAGATACAGAGACGCGCTCATCTTTGGAT 3600  
5987 GCTTCCCAATGCGCATGAAAGCGGCCAGAGATACAGAGACGCGCTCATCTTTGGAT 6046  
3601 CCGGATCTGCTTCTTACTACGATCAGCTTCTTCAAGGAAATGAGCAAGTGGCATGCTC 3660  
6047 CCGGATCTGCTTCTTACTACGATCAGCTTCTTCAAGGAAATGAGCAAGTGGCATGCTC 6106  
3661 GCACTGCTATTCAGACTCTGAAGAACCGTATCGGCAAGTGCAGAAATATGCGAG 3720  
6107 GCACTGCTATTCAGACTCTGAAGAACCGTATCGGCAAGTGCAGAAATATGCGAG 6166  
3721 TTGGTCAAAACAGAGATGAGAGACAGCAAGAACCTTACCTTCCGCTCAACAG 3780  
6167 TTGGTCAAAACAGAGATGAGAGACAGCAAGAACCTTACCTTCCGCTCAACAG 6226  
3781 GTTTATGAAAATGTGGGCTGATCAGGCTGAGGCGATGAGCAAAATCGGAGCAATTTAT 3840  
6227 GTTTATGAAAATGTGGGCTGATCAGGCTGAGGCGATGAGCAAAATCGGAGCAATTTAT 6286  
3841 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGAACCTGATGATGATACA 3900  
6287 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGAACCTGATGATGATACA 6346  
3901 TGGGCAATGCTGAGGCTTACAGGCTTTTCAAGCTGTATACCAAGAGCCCAAGTTC 3960  
6347 TGGGCAATGCTGAGGCTTACAGGCTTTTCAAGCTGTATACCAAGAGCCCAAGTTC 6406  
3961 GTTGGGAGATGGCGGCGACAGCTGCGTACATTTAAGGCGAGATGAGAGCAGACC 4020  
6407 GTTGGGAGATGGCGGCGACAGCTGCGTACATTTAAGGCGAGATGAGAGCAGACC 6466  
4021 GGTGAAGGCGCGCGGCTTATGACCGGCTTCAATGATGCGGCTTATGCGGATAG 4080  
6467 GGTGAAGGCGCGCGGCTTATGACCGGCTTCAATGATGCGGCTTATGCGGATAG 6526  
4081 AAGTTTACGAAGCAGTATGTGGCCAGGCTGAGAGCGATGAGATCGTACCTGATCCG 4140  
6527 AAGTTTACGAAGCAGTATGTGGCCAGGCTGAGAGCGATGAGATCGTACCTGATCCG 6586  
4141 GAGGCTATGAGTGTGGGCGAGATGATTTTGAATGGAATTTGTTTCAAGGGAATGGC 4200  
6587 GAGGCTATGAGTGTGGGCGAGATGATTTTGAATGGAATTTGTTTCAAGGGAATGGC 6646  
4201 GATTAT 4206

|||||  
Db 6647 GATTAT 6652

RESULT 2  
ADN40832  
ID ADN40832 standard; DNA; 4206 BP.  
XX  
AC ADN40832;  
XX  
XX 15-JUL-2004 (first entry)  
DT  
XX  
DE Neurospora crassa QDE-1 (quelling defective) DNA SeqID 1.  
XX  
XX soluble RNA polymerase; post-transcriptional gene silencing; PTGS;  
KM microarray; RNA interference; RNA binding protein; QDE-1;  
XX quelling defective; gene; ds.  
OS Neurospora crassa.  
XX  
XX Key Location/Qualifiers  
FH 1. 4206  
FT CDS /\*tag= a  
FT /product= "QDE-1 protein"  
XX  
XX MO2004035784-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 17-OCT-2003; 2003WO-F1000776.  
XX  
XX 21-OCT-2002; 2002US-0419562P.  
XX  
XX (RNAL-) RNA LINE OY.  
XX  
XX Mekeyev E, Bamford D;  
XX  
XX WPI; 2004-399869/37.  
XX  
XX P-PSDB; ADN40833.  
XX  
XX  
XX Producing nucleic acid product, involves contacting with template an RNA  
PT polymerase which produces short complementary RNA copies of template,  
PT that are scattered throughout entire template, and template-length  
PT complementary RNA copies.  
XX  
XX  
XX Claim 11; SEQ ID NO 1; 95bp; English.  
XX  
XX  
XX This invention relates to a novel method for producing a soluble RNA  
CC polymerase protein. Specifically, it refers to an RNA polymerase enzyme  
CC isolated from the post-transcriptional gene silencing (PTGS) pathway that  
CC is useful for in vitro RNA synthesis kits. The present invention  
CC describes an enzyme with enhanced solubility capable of producing two  
CC types of reaction products i.e. both short and long RNA copies of the  
CC template including template length complementary RNA copies. Furthermore,  
CC it can also copy single stranded DNA templates in reactions that do not  
CC require a primer for initiation of RNA synthesis. As such, this method  
CC can also be used to produce radioactively or chemically labelled RNA  
CC probes for in situ hybridisation analyses or microarray work, to generate  
CC trigger RNA molecules to induce RNA interference effects in vivo and in  
CC vitro or to study nucleic acid secondary structure, nucleic acid-protein  
CC interactions or more preferably RNA-protein interactions. This  
CC polynucleotide sequence is Neurospora crassa DNA that encodes the RNA  
CC binding protein identified as QDE-1 (quelling defective), given in an  
CC exemplification of the invention.  
XX  
XX  
XX Sequence 4206 BP; 1038 A; 1141 C; 1113 G; 914 T; 0 U; 0 Other;  
SQ

Query Match 99.9%; Score 4202.8; DB 12; Length 4206;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCTTATTCTCTAGAGAGAGAAATGACCCCTCGAGAAATCATTAACCGGCTC 60  
|||||

Db 1 ATGAACCTTATTCTCTAGAGAGAGAAATGACCCCTCGAGAAATCATTAACCGGCTC 60  
Qy  
Qy 61 AATAACGACTTACCACTGGGCTCTCAGTGTGTGACAGACAACTCTCACCCCCACGG 120  
Db 61 AATAACGACTTACCACTGGGCTCTCAGTGTGTGACAGACAACTCTCACCCCCACGG 120  
Qy 121 CGGAGAGCTGGCCGAGAGTGAAGAGATTTTCGGTCCGCATGACAAGATCTACAGAGCC 180  
Db 121 CGGAGAGCTGGCCGAGAGTGAAGAGATTTTCGGTCCGCATGACAAGATCTACAGAGCC 180  
Qy 181 CTGAACCTTTCTCTACTGGCGGAGAGATGATCCCTGTAACCAAGGCAAGCACTTCTTC 240  
Db 181 CTGAACCTTTCTCTACTGGCGGAGAGATGATCCCTGTAACCAAGGCAAGCACTTCTTC 240  
Qy 241 ATGAGAGCCAAAGCTCGAGCTGAACTGAGTGGTCCCAAGCCGACCTGACAGC 300  
Db 241 ATGAGAGCCAAAGCTCGAGCTGAACTGAGTGGTCCCAAGCCGACCTGACAGC 300  
Qy 301 CTTCCGTGTGTCAGAGAACTCCCGCGCGCTTACTGCGGCGCAACATGGGCAATTGCA 360  
Db 301 CTTCCGTGTGTCAGAGAACTCCCGCGCGCTTACTGCGGCGCAACATGGGCAATTGCA 360  
Qy 361 ACTGTGTTGCTGAGGTGCTTAATAGGTTATGCACTCCCAATTAACAACAGGATCGA 420  
Db 361 ACTGTGTTGCTGAGGTGCTTAATAGGTTATGCACTCCCAATTAACAACAGGATCGA 420  
Qy 421 ACCTTTGGCAGAACTTAAGCGGCGCCAGTGGCGCTGAGCGGCCCAACTCTACCAACCC 480  
Db 421 ACCTTTGGCAGAACTTAAGCGGCGCCAGTGGCGCTGAGCGGCCCAACTCTACCAACCC 480  
Qy 481 AAACGAGAGATGAGCCCGCAATGTCACTTTGCTGATCCGCCCAACGCTGTTGACT 540  
Db 481 AAACGAGAGATGAGCCCGCAATGTCACTTTGCTGATCCGCCCAACGCTGTTGACT 540  
Qy 541 CGCTTGCCACAGAGTCTCTATTCAAGGCGGCGGATACCCCTTAAGTTCCCGGATCA 600  
Db 541 CGCTTGCCACAGAGTCTCTATTCAAGGCGGCGGATACCCCTTAAGTTCCCGGATCA 600  
Qy 601 GTGAATACCGGTTTCAAGAGCAACCATCTCTGAGAGTGAATCTCAATGATGACCAAG 660  
Db 601 GTGAATACCGGTTTCAAGAGCAACCATCTCTGAGAGTGAATCTCAATGATGACCAAG 660  
Qy 661 CGGCGCAAGGGCAAGCTGTCTGATTAATGTTGCCGTGCGCGCCGCCGCTGCTATT 720  
Db 661 CGGCGCAAGGGCAAGCTGTCTGATTAATGTTGCCGTGCGCGCCGCCGCTGCTATT 720  
Qy 721 GCGAGGCTTTGGACAGGATCCGACTCGAAGGCAATGCAATGAGAGATGCCAGGGG 780  
Db 721 GCGAGGCTTTGGACAGGATCCGACTCGAAGGCAATGCAATGAGAGATGCCAGGGG 780  
Qy 781 ACAAGTCATAGACGAGCGGACAGAGTGAATTCCTTGTATATCTTCAGGCACTTCTAT 840  
Db 781 ACAAGTCATAGACGAGCGGACAGAGTGAATTCCTTGTATATCTTCAGGCACTTCTAT 840  
Qy 841 GGTTCAGAGTCTTTCAGCGCTTGGCGCTTCAATCAGAGCACTTACCAAGTATTTTGA 900  
Db 841 GGTTCAGAGTCTTTCAGCGCTTGGCGCTTCAATCAGAGCACTTACCAAGTATTTTGA 900  
Qy 901 GCTTCCTCTTCAAGCGCAGAGAGAGCGGCTGTGATGCAAGCGCTTTGAGGTGGA 960  
Db 901 GCTTCCTCTTCAAGCGCAGAGAGAGCGGCTGTGATGCAAGCGCTTTGAGGTGGA 960  
Qy 961 CACTTATGAGTCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Db 961 CACTTATGAGTCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Qy 1021 CTTTCATGCTCTTCCAGGAGTGAATCTGTTGAGCACTTACTATGAGTGTTCAGT 1080  
Db 1021 CTTTCATGCTCTTCCAGGAGTGAATCTGTTGAGCACTTACTATGAGTGTTCAGT 1080  
Qy 1081 TCCGCGCGGAGGAGGAGCAATTCGCGAGCGGAGTCCGTAATGAGACTGGCTGGAGCGAA 1140  
Db 1081 TCCGCGCGGAGGAGGAGCAATTCGCGAGCGGAGTCCGTAATGAGACTGGCTGGAGCGAA 1140

QY 1141 GAAAGCGTGCATCTCAGGTTCAAGTTCAATGCCCGGATGTTGAGCTCGGCTGAGAAAT 1200  
DB 1141 GAAAGCGCTGCATCTCAGGTTCAAGTTCAATGCCCGGATGTTGAGCTCGGCTGAGAAAT 1200  
QY 1201 ATTTGGCGAAATTTCCCAATGGCTACAGAACTCCTCTCGTGTTCATGGAAATT 1260  
DB 1201 ATTTGGCGAAATTTCCCAATGGCTACAGAACTCCTCTCGTGTTCATGGAAATT 1260  
QY 1261 ACCGACTCTTTATGCACTGCAAGTAGACTTTGGAAGCAGAGACCTGGGCTTAAGTAC 1320  
DB 1261 ACCGACTCTTTATGCACTGCAAGTAGACTTTGGAAGCAGAGACCTGGGCTTAAGTAC 1320  
QY 1321 GACCTTCTGCTGCTACCGGCGGATGTCAGATATCTGGAAGACTCTACCGGCTT 1380  
DB 1321 GACCTTCTGCTGCTACCGGCGGATGTCAGATATCTGGAAGACTCTCTACCGGCTT 1380  
QY 1381 GATGCTTTCCGTGGTAAACCTTTTCAGAAAAGCCGCCAAGACGTGTGTAACGCA 1440  
DB 1381 GATGCTTTCCGTGGTAAACCTTTTCAGAAAAGCCGCCAAGACGTGTGTAACGCA 1440  
QY 1441 ATGACGGGCAACTTTGAGAGCAAAAGTAGTCCGTTGTTCTCTGCTGTTCTAGACTAC 1500  
DB 1441 ATGACGGGCAACTTTGAGAGCAAAAGTAGTCCGTTGTTCTCTGCTGTTCTAGACTAC 1500  
QY 1501 AATCGGCAACTGCGCTACTGCGGCCCTTTACCTTGGAAGCTGGAAGCGCTGATGTC 1560  
DB 1501 AATCGGCAACTGCGCTACTGCGGCCCTTTACCTTGGAAGCTGGAAGCGCTGATGTC 1560  
QY 1561 GAGCAGGCTGTGCACTGCAACCGCTCGGTTCCGCTCTGATAGTTTTCAGATCTTATA 1620  
DB 1561 GAGCAGGCTGTGCACTGCAACCGCTCGGTTCCGCTCTGATAGTTTTCAGATCTTATA 1620  
QY 1621 CCGTGCCTACGAGACCAAGCCCAAGTGTACCGCGGTGTGACAAACAACAGGTGCG 1680  
DB 1621 CCGTGCCTACGAGACCAAGCCCAAGTGTACCGCGGTGTGACAAACAACAGGTGCG 1680  
QY 1681 GTGCAAGAGTCACTGCACTGCTGCAAGTGGGCAACCTCTGTTGAGGCGGCAATGG 1740  
DB 1681 GTGCAAGAGTCACTGCACTGCTGCAAGTGGGCAACCTCTGTTGAGGCGGCAATGG 1740  
QY 1741 CGGCTTTCTTCCGCAAAAGATGCGGATACAGAAACCTCTCAAGGAGTTCAAGTCCG 1800  
DB 1741 CGGCTTTCTTCCGCAAAAGATGCGGATACAGAAACCTCTCAAGGAGTTCAAGTCCG 1800  
QY 1801 GCCGAGACCCGAAACCCATATCAAGAGAGAGTCCACTTCTTTGCGGAGACCGGCAAT 1860  
DB 1801 GCCGAGACCCGAAACCCATATCAAGAGAGAGTCCACTTCTTTGCGGAGACCGGCAAT 1860  
QY 1861 ACCTTCCGACTGATGTGTTCAAGACGAGATCTGTCTGCGGAGAGAACTGTAGAG 1920  
DB 1861 ACCTTCCGACTGATGTGTTCAAGACGAGATCTGTCTGCGGAGAGAACTGTAGAG 1920  
QY 1921 CAAGCAGACCGAGTTCAAGTTAGTCAAAATGCTGAGCTGCTCTGCACTGCAAGAAC 1980  
DB 1921 CAAGCAGACCGAGTTCAAGTTAGTCAAAATGCTGAGCTGCTCTGCACTGCAAGAAC 1980  
QY 1981 ACTTGGCAGCGCACTCAAGTTGTTCTCCGATACAGCTCGGTCTGATGATAGACATAT 2040  
DB 1981 ACTTGGCAGCGCACTCAAGTTGTTCTCCGATACAGCTCGGTCTGATGATAGACATAT 2040  
QY 2041 GCCATTATGACATTTGAGAGCTTCAAGATCAAGACCAAGACCGATCTTTCAACT 2100  
DB 2041 GCCATTATGACATTTGAGAGCTTCAAGATCAAGACCAAGACCGATCTTTCAACT 2100  
QY 2101 TCAAGCAGTGGCAGAGTATGATGATGAGCTGATGAGCCGATGTGCGAAGGTGGCCAG 2160  
DB 2101 TCAAGCAGTGGCAGAGTATGATGATGAGCTGATGAGCCGATGTGCGAAGGTGGCCAG 2160  
QY 2161 AGGATACCGCATGTTCTCGGTTTGGGATGATGCTCTGCTGTCAGAGGCGGTTTGT 2220  
DB 2161 AGGATACCGCATGTTCTCGGTTTGGGATGATGCTCTGCTGTCAGAGGCGGTTTGT 2220

QY 2221 TCGGCCAAGGAAATGCTGATATGACGTTGACGACACAGGCGATGAGATTGATGAG 2280  
DB 2221 TCGGCCAAGGAAATGCTGATATGACGTTGACGACACAGGCGATGAGATTGATGAG 2280  
QY 2281 ACATACCCGTTCCAGGCGCAAGTGGAAATGCACTTGTGATTAACATCAAGTACCTC 2340  
DB 2281 ACATACCCGTTCCAGGCGCAAGTGGAAATGCACTTGTGATTAACATCAAGTACCTC 2340  
QY 2341 GAATCCGAGCGTGGCTTCTGAACTGAACTGAGCTGCTCAACCTACAGCTGTTACT 2400  
DB 2341 GAATCCGAGCGTGGCTTCTGAACTGAACTGAGCTGCTCAACCTACAGCTGTTACT 2400  
QY 2401 GTCCTGAAATATGAGCAGGAGCAAGAGTGAATGCGCAGCAATCGGTGACCGCTT 2460  
DB 2401 GTCCTGAAATATGAGCAGGAGCAAGAGTGAATGCGCAGCAATCGGTGACCGCTT 2460  
QY 2461 ATCAACGATTTGCAACGACAGTTTACGAGCAGCAAAAGCATCTTTGAATCCGCCAGTGGAA 2520  
DB 2461 ATCAACGATTTGCAACGACAGTTTACGAGCAGCAAAAGCATCTTTGAATCCGCCAGTGGAA 2520  
QY 2521 TTTTCCCAATGGGTTTACGAGATTAATTCAGTCCGCACTGAGTCAGCAGCGCGT 2580  
DB 2521 TTTTCCCAATGGGTTTACGAGATTAATTCAGTCCGCACTGAGTCAGCAGCGCGT 2580  
QY 2581 GTGCTTTCTTCTGCTGCGCTACCTGACAGTCAAGAGAGACACTGAACCTTTGATGAC 2640  
DB 2581 GTGCTTTCTTCTGCTGCGCTACCTGACAGTCAAGAGAGACACTGAACCTTTGATGAC 2640  
QY 2641 AGTGGGTTGATCCCAAGAAAGCAAAAGTACTTGCAGACATCCGCTGGATCTTCAAAAG 2700  
DB 2641 AGTGGGTTGATCCCAAGAAAGCAAAAGTACTTGCAGACATCCGCTGGATCTTCAAAAG 2700  
QY 2701 CGGAAATGTAACAGTTGAAGTCAAGCTGAACATCCGCTGCTGATGATGATGAT 2760  
DB 2701 CGGAAATGTAACAGTTGAAGTCAAGCTGAACATCCGCTGCTGATGATGATGATGAT 2760  
QY 2761 TACATGATTCGCAATTTCTGCGGTGCTTGAAGAAATGAGTTCAATGTCGATCTCC 2820  
DB 2761 TACATGATTCGCAATTTCTGCGGTGCTTGAAGAAATGAGTTCAATGTCGATCTCC 2820  
QY 2821 TCAAGTTTCAAGAGCAGAGAGAGTCTTTTACACTCTTATGCACTGTATGTCCTG 2880  
DB 2821 TCAAGTTTCAAGAGCAGAGAGAGTCTTTTACACTCTTATGCACTGTATGTCCTG 2880  
QY 2881 GCGGATCCCCCAGGCCATTTCCCTAGATATCAAGGTTGAGAGAGCTTCAAGCA 2940  
DB 2881 GCGGATCCCCCAGGCCATTTCCCTAGATATCAAGGTTGAGAGAGCTTCAAGCA 2940  
QY 2941 GAGCTCCAGAGTCAAGAGTGTATCATCTTCTACTAAGAGATGTACCGCTTGTCT 3000  
DB 2941 GAGCTCCAGAGTCAAGAGTGTATCATCTTCTACTAAGAGATGTACCGCTTGTCT 3000  
QY 3001 AAGAGCTATCTGCTGAGAGCTACAGACGCGATATGAGCTGCTGCTGGATCCGAG 3060  
DB 3001 AAGAGCTATCTGCTGAGAGCTACAGACGCGATATGAGCTGCTGCTGGATCCGAG 3060  
QY 3061 ATGCTCATGTTCTGCTCAATGCGGAAATGCTTGTGAAGCCCACTGTCTAGGTACTTA 3120  
DB 3061 ATGCTCATGTTCTGCTCAATGCGGAAATGCTTGTGAAGCCCACTGTCTAGGTACTTA 3120  
QY 3121 AAGAGGACAAAAGCACTTTCAAAACAATTATGCTCACAAGGCAAGGCTCAGCGGCC 3180  
DB 3121 AAGAGGACAAAAGCACTTTCAAAACAATTATGCTCACAAGGCAAGGCTCAGCGGCC 3180  
QY 3181 AAGAGCAGACTACATACATATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
DB 3181 AAGAGCAGACTACATACATATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
QY 3241 TTTCTTGGCAGTGTGCTATCACTAAGAAAGAGTCTGTTACATCAATAGTGTCT 3300  
DB 3241 TTTCTTGGCAGTGTGCTATCACTAAGAAAGAGTCTGTTACATCAATAGTGTCT 3300  
QY 3301 AACAAAGCGGCAATCTTCTAGTTCACTGCTGGGAAACCTGTCATGAGAGCAAGCA 3360

```

Dh 3301 AACAGCCGGGCATCTTCTAGTCACTGATGGAAACCTCGTCAATCAGAGAAACAA
Qy 3361 GGTATGCTCTTTAACAAGCAAGCTGGGCTCAATTGGTGAAGGAACGTTGGCGGTGCA
Dh 3361 GGTATGCTCTTTAACAAGCAAGCTGGGCTCAANTGGTGAAGGAACGTTGGCGGTGCA
Qy 3421 TTGTCCCTCTCTGACCAATGATACAGAGCAAGTGGCTCGGGCCGCGAGACCTACC
Dh 3421 TTGTCCCTCTCTGACCAATGATACAGAGCAAGTGGCTCGGGCCGCGAGACCTACC
Qy 3481 CACATTTATGACTACCTGAAATCTCGATGCCAGGCTCGATGTAACAAGAACTGGAA
Dh 3481 CACATTTATGACTACCTGAAATCTCGATGCCAGGCTCGATGTAACAAGAACTGGAA
Qy 3541 GCCTTCACAATGCAATGAAAGCGCCAGAGATACAGAAACGCGCTCACTTTGGGAT
Dh 3541 GCCTTCACAATGCAATGAAAGCGCCAGAGATACAGAAACGCGCTCACTTTGGGAT
Qy 3601 CCGGATCTCGCTTCTTAATAAGCTTTTCAAGAGATTAGCAAGATGCGCATGCTCC
Dh 3601 CCGGATCTCGCTTCTTAATAAGCTTTTCAAGAGATTAGCAAGATGCGCATGCTCC
Qy 3661 GCACTGCTATTCAAGACTGAGAGAACCCGATCGGCGAATTCAGAGAAAGAAATGGAGG
Dh 3661 GCACTGCTATTCAAGACTGAGAGAACCCGATCGGCGAATTCAGAGAAAGAAATGGAGG
Qy 3721 TTGGTCAAAAACAAGAGATGAGAGACAGCAAGAACCCCTACCTGTCGCCGTCACACAG
Dh 3721 TTGGTCAAAAACAAGAGATGAGAGACAGCAAGAACCCCTACCTGTCGCCGTCACACAG
Qy 3781 GTTTATGAAAAATGATGCGCCATCAAGCTGAGAGCGATGCAAAATCCGAGCAAAATTAT
Dh 3781 GTTTATGAAAAATGATGCGCCATCAAGCTGAGAGCGATGCAAAATCCGAGCAAAATTAT
Qy 3841 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCCTCTCGCGGACCGTGAAGAAATACA
Dh 3841 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCCTCTCGCGGACCGTGAAGAAATACA
Qy 3901 TGGGCAATGCTGAGGCTGACAGGCTTTCAAGCTGTAACAACAAGAACCCCAAGTTT
Dh 3901 TGGGCAATGCTGAGGCTGACAGGCTTTCAAGCTGTAACAACAAGAACCCCAAGTTT
Qy 3961 GTGTGCAAGATGCGGCGCAACAGCTTGCCTAATTAAGCGCAGATGACAGCAAGCCC
Dh 3961 GTGTGCAAGATGCGGCGCAACAGCTTGCCTAATTAAGCGCAGATGACAGCAAGCCC
Qy 4021 GGTGAAGGCGCCCGGGCTGATGACCGGCTGATGATGCGGCTTGAATGCGCGATAG
Dh 4021 GGTGAAGGCGCCCGGGCTGATGACCGGCTGATGATGCGGCTTGAATGCGCGATAG
Qy 4081 AAGTTTACGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Dh 4081 AAGTTTACGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Qy 4141 GAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Dh 4141 GAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Qy 4201 GATTAT 4206
Dh 4201 GATTAT 4206

```

```

XX KW soluble RNA polymerase; post-transcriptional gene silencing; PTGS;
KM microarray; RNA interference; RNA binding protein; QDE-1;
XX KM quelling defective; gene; ds.
OS Neurospora crassa.
XX FH Key Location/Qualifiers
FT CDS 1..3078
FT FT /*tag= a
FT FT /partial
FT FT /product= "truncated QDE-1 protein"
FT FT /note= "Start and stop codons are absent"
XX PN MO2004035784-A2.
XX 29-APR-2004.
XX 17-OCT-2003; 2003MO-F1000776.
XX 21-OCT-2002; 2002US-0419562P.
XX PA (RNAL-) RNA LINE OY.
XX MEkeyev E, Bamford D;
XX WPI; 2004-399869/37.
XX P-PSDB; ADN40835.
XX PT Producing nucleic acid product, involves contacting with template an RNA
PT polymerase which produces short complementary RNA copies of template,
PT that are scattered throughout entire template, and template-length
PT complementary RNA copies.
XX PS Claim 11; SEQ ID NO 3; 95pp; English.
XX CC This invention relates to a novel method for producing a soluble RNA
XX CC polymerase protein. Specifically, it refers to an RNA polymerase enzyme
XX CC isolated from the post-transcriptional gene silencing (PTGS) pathway that
XX CC is useful for in vitro RNA synthesis kits. The present invention
XX CC describes an enzyme with enhanced solubility capable of producing two
XX CC types of reaction products i.e. both short and long RNA copies of the
XX CC template including template length complementary RNA copies. Furthermore,
XX CC it can also copy single stranded DNA templates in reactions that do not
XX CC require a primer for initiation of RNA synthesis. As such, this method
XX CC can also be used to produce radioactively or chemically labelled RNA
XX CC probes for in situ hybridisation analyses or microarray work, to generate
XX CC trigger RNA molecules to induce RNA interference effects in vivo and in
XX CC vitro or to study nucleic acid secondary structure, nucleic acid-protein
XX CC interactions or more preferably RNA-protein interactions. This
XX CC polynucleotide is a truncated Neurospora crassa DNA sequence that encodes
XX CC the RNA binding protein identified as QDE-1 (quelling defective), given
XX CC in an exemplification of the invention.
SQ Sequence 3078 BP; 763 A; 785 C; 835 G; 695 T; 0 U; 0 Other;
Query Match 73.2%; Score 3078; DB 12; Length 3078;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3078; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1129 GCTGGAGCGAAGAAAGCGCTGATCTCAGTTCAAGTTCAATGCCCCGGTGGTGACGT 1188
Dh 1 GCTCGAGCGAAGAAAGCGCTGATCTCAGTTCAAGTTCAATGCCCCGGTGGTGACGT 60
Qy 1189 CGGCTGAGAAATATTTGGCGGAATTTCCCAATGCTACAGAAAGCTCTCTCGGTGT 1248
Dh 61 CGGCTGAGAAATATTTGGCGGAATTTCCCAATGCTACAGAAAGCTCTCTCGGTGT 120
Qy 1249 GCATGGAGATTACAGACTCTTTATGACTGCAAGTGAAGTTGAGAGACGAGACCTG 1308
Dh 121 GCATGGAGATTACAGACTCTTTATGACTGCAAGTGAAGTTGAGAGACGAGACCTG 180
Qy 1309 GGCCTAAAGTACGACCTTCTGTTCTACCGGCGCGATGTGACAGATATCTGGAAGACT 1368

```

Db 181 GGCCTTAAGTACGACCTCTCTGCTCTACCGCGGCGATGTCACAGATATCTGGAAGACT 240  
Qy 1369 CTCTACCGGCTTGTATGCTCTTCCGTGTAAACCTTTTCCAGAAAACCGCCCAACGACGTG 1428  
Db 241 CTCTACCGGCTTGTATGCTCTTCCGTGTAAACCTTTTCCAGAAAACCGCCCAACGACGTG 300  
Qy 1429 TTGCGTACCGCAATGAGCGGGCACTTTGAGGCAAAAGTAGTGGCGTGTCTCTGCT 1488  
Db 301 TTGCGTACCGCAATGAGCGGGCACTTTGAGGCAAAAGTAGTGGCGTGTCTCTGCT 360  
Qy 1489 GTTCTAGACTACAAATCCGCAACCTCGCTACCTGCGCCCTTTTACTTGTGAAGCTGAAG 1548  
Db 361 GTTCTAGACTACAAATCCGCAACCTCGCTACCTGCGCCCTTTTACTTGTGAAGCTGAAG 420  
Qy 1549 CCGCTCATGTTTGAAGAGGGCTGTGACTCACCCGTGCTTCCGTCTGATAGGTTTTTC 1608  
Db 421 CCGCTCATGTTTGAAGAGGGCTGTGACTCACCCGTGCTTCCGTCTGATAGGTTTTTC 480  
Qy 1609 GAGATCCCTTATACCGTGGCCCTAGAGACCAAGCCCAAGTGTACCGCGGTGGTCAAGAAA 1668  
Db 481 GAGATCCCTTATACCGTGGCCCTAGAGACCAAGCCCAAGTGTACCGCGGTGGTCAAGAAA 540  
Qy 1669 CAACCAAGTGGCTGCAAGAGATCATCCAGTGCCTCAAGATGGGGCAACATCTCTTGTA 1728  
Db 541 CAACCAAGTGGCTGCAAGAGATCATCCAGTGCCTCAAGATGGGGCAACATCTCTTGTA 600  
Qy 1729 GCGCGCAATGGCGCGCTTTCTTGCCAAAGATCCCGGATACAGAAAACCTCTCAGGAG 1788  
Db 601 GCGCGCAATGGCGCGCTTTCTTGCCAAAGATCCCGGATACAGAAAACCTCTCAGGAG 660  
Qy 1789 TTCCAGCTCGCGCGCGAGAACCCGAACCCATCATCAAGAGAGAGTCCACTTTTGCC 1848  
Db 661 TTCCAGCTCGCGCGCGAGAACCCGAACCCATCATCAAGAGAGAGTCCACTTTTGCC 720  
Qy 1849 GAGACCGGCAATTAAGCTTCCGACTGATGTGTTCAAGAAGAGATGTGTTCCGGCAGAG 1908  
Db 721 GAGACCGGCAATTAAGCTTCCGACTGATGTGTTCAAGAAGAGATGTGTTCCGGCAGAG 780  
Qy 1909 GAACCTGTAGAGCAACGAGCCGAGTTCAAGTTAGTCAATGTGAGTGGCTCTTGCA 1968  
Db 781 GAACCTGTAGAGCAACGAGCCGAGTTCAAGTTAGTCAATGTGAGTGGCTCTTGCA 840  
Qy 1969 CTGCAACAACAACCTTGGCAGCGGCACTCAAGTTTCTCCGATCCAGCTGGCTG 2028  
Db 841 CTGCAACAACAACCTTGGCAGCGGCACTCAAGTTTCTCCGATCCAGCTGGCTG 900  
Qy 2029 AGTAAACATATGCAATTTATGACATTTGAGGCTCAGCAGATCAGACACACAAGCCGAT 2088  
Db 901 AGTAAACATATGCAATTTATGACATTTGAGGCTCAGCAGATCAGACACACAAGCCGAT 960  
Qy 2089 CTTCTTTTCACTTCAAGGCACTGCGCAAGTGAATGAATGCGGTGAGGCCGATCTCGGA 2148  
Db 961 CTTCTTTTCACTTCAAGGCACTGCGCAAGTGAATGAATGCGGTGAGGCCGATCTCGGA 1020  
Qy 2149 AGCGTGGCAAGAGATACCGGATGTTCTCGTTTGGTGAATGGCCCTGCTGTGAA 2208  
Db 1021 AGCGTGGCAAGAGATACCGGATGTTCTCGTTTGGTGAATGGCCCTGCTGTGAA 1080  
Qy 2209 GGGCGGTTTGGTTGGGCAAGGAGATGGGTTATCCAGCTTGAAGACACAGCGGATGAG 2268  
Db 1081 GGGCGGTTTGGTTGGGCAAGGAGATGGGTTATCCAGCTTGAAGACACAGCGGATGAG 1140  
Qy 2269 GATTGATCGAGACATACCGCTCCAGCGCAAGTGGAGATCGACTTGTGTTGAATTAACAT 2328  
Db 1141 GATTGATCGAGACATACCGCTCCAGCGCAAGTGGAGATCGACTTGTGTTGAATTAACAT 1200  
Qy 2329 CAAGTATCCCTGGAAGTCCGAGGGCTTCTGAATCGAAGTGAAGTGGCTGTAAACCTA 2388  
Db 1201 CAAGTATCCCTGGAAGTCCGAGGGCTTCTGAATCGAAGTGAAGTGGCTGTAAACCTA 1260  
Qy 2389 CAGCTTTACTGTCTGGAAGATAGAGCCAGGACAAAGTGAAGATGCCCAGGCAATC 2448

Db 1261 CAGCTTTACTGTCTGGAAGATAGAGCCAGGGAACAAGTGAAGATGCCACAGCAATC 1320  
Qy 2449 GGTGACCGCTTTATCAACGATTTTGCAACGACAGTTCAAGCGAAGAAAAGCATGCTTTGAAT 2508  
Db 1321 GGTGACCGCTTTATCAACGATTTTGCAACGACAGTTCAAGCGAAGAAAAGCATGCTTTGAAT 1380  
Qy 2509 CGCCAGTGAATTTGCGCAATGGGTTTACGAGATTAATCCAGTGGCGCACTGAGTC 2568  
Db 1381 CGCCAGTGAATTTGCGCAATGGGTTTACGAGATTAATCCAGTGGCGCACTGAGTC 1440  
Qy 2569 AGCCACGGCCGTGTCCTTTCTTGTGGCTACCTGACAGTCAAGAGAGACACTGAAC 2628  
Db 1441 AGCCACGGCCGTGTCCTTTCTTGTGGCTACCTGACAGTCAAGAGAGACACTGAAC 1500  
Qy 2629 TTCTGTGAACAGTGGGTTCCATCCCAAGAACCAAGTAACTTTGCAAGACATGCGCTGG 2688  
Db 1501 TTCTGTGAACAGTGGGTTCCATCCCAAGAACCAAGTAACTTTGCAAGACATGCGCTGG 1560  
Qy 2689 GATCTTCAAAAGCGGAAATGTGACAGTTGAAGTCCAAAGCTGAACATCCGTGCGTCA 2748  
Db 1561 GATCTTCAAAAGCGGAAATGTGACAGTTGAAGTCCAAAGCTGAACATCCGTGCGTCA 1620  
Qy 2749 TCAAGATACATTTATCATGATTTGCGGCTGTGCTGAGAGAAAATGAGTTTCAAT 2808  
Db 1621 TCAAGATACATTTATCATGATTTGCGGCTGTGCTGAGAGAAAATGAGTTTCAAT 1680  
Qy 2809 GTGCGATTTCTCTCAAAAGTTCAAGACAGAGAGAGAGCTTTTACATCTCTATCGGACTGT 2868  
Db 1681 GTGCGATTTCTCTCAAAAGTTCAAGACAGAGAGAGAGCTTTTACATCTCTATCGGACTGT 1740  
Qy 2869 GATGTCCTGCTGGGCGGATCCCGAGCCCATTTCCCTAGTATATCCAAAGGTTGAGCA 2928  
Db 1741 GATGTCCTGCTGGGCGGATCCCGAGCCCATTTCCCTAGTATATCCAAAGGTTGAGCA 1800  
Qy 2929 GTCTTCAAGCAGAGCTCAAGTCTCAAGATGTATCATCTTCTACTAAAGAGAT 2988  
Db 1801 GTCTTCAAGCAGAGCTCAAGTCTCAAGATGTATCATCTTCTACTAAAGAGAT 1860  
Qy 2989 GTACCGTTTGTAAAGACTATCTGGTGAAGACTACACGCGATATGAGCTGGCTGTGC 3048  
Db 1861 GTACCGTTTGTAAAGACTATCTGGTGAAGACTACACGCGATATGAGCTGGCTGTGC 1920  
Qy 3049 TGGGATCCGAGATTCGTGATGTTTCGTCAATGCGGAAATGCTCGAGCCGACCTG 3108  
Db 1921 TGGGATCCGAGATTCGTGATGTTTCGTCAATGCGGAAATGCTCGAGCCGACCTG 1980  
Qy 3109 TCTAGTACTTAAAGAGAGCAAAACGACTTTCAACCACTTATGCGCTCAACGAGCAG 3168  
Db 1981 TCTAGTACTTAAAGAGAGCAAAACGACTTTCAACCACTTATGCGCTCAACGAGCAG 2040  
Qy 3169 GGTCAAGCGCCAAAGAGAGACTACATACGATATGATCCAGAAAGCTTCCATTTGCGC 3228  
Db 2041 GGTCAAGCGCCAAAGAGAGACTACATACGATATGATCCAGAAAGCTTCCATTTGCGC 2100  
Qy 3229 CTGAGGCCCAACTTCTTGGGCAATGTGCACTAACCAAGAAAGGCTCTGTACATCAAC 3288  
Db 2101 CTGAGGCCCAACTTCTTGGGCAATGTGCACTAACCAAGAAAGGCTCTGTACATCAAC 2160  
Qy 3289 AATAGTGTCTTAAACAAGCCGACATCATTTAGTTCACTGGTGGGAAAACCTGTGAT 3348  
Db 2161 AATAGTGTCTTAAACAAGCCGACATCATTTAGTTCACTGGTGGGAAAACCTGTGAT 2220  
Qy 3349 CAGAGCAAGCAAGTATTTGCTTTTAAAGCAAGCAAGCTGGGCTCAATTTGCTAAGGAATG 3408  
Db 2221 CAGAGCAAGCAAGTATTTGCTTTTAAAGCAAGCAAGCTGGGCTCAATTTGCTAAGGAATG 2280  
Qy 3409 CTTGGCGGTGATTTGCTTCTCTGACCCCAATGTACAAAGAGCAAGTGGCTCGGGCGC 3468  
Db 2281 CTTGGCGGTGATTTGCTTCTCTGACCCCAATGTACAAAGAGCAAGTGGCTCGGGCGC 2340  
Qy 3469 GAGAGGCTTACCAACATTAATGACTACCTGAAAATCTGCATGCGCAAGGCTGCGATTTGAC 3528  
Db 2341 GAGAGGCTTACCAACATTAATGACTACCTGAAAATCTGCATGCGCAAGGCTGCGATTTGAC 2400

```

QY 3529 AAGGAATGGAAGCTTCCCAATGCCATGAAAGCGGCCAAGATACAGAAGACGCGCT 3588
D 2401 AAGGAATGGAAGCTTCCCAATGCCATGAAAGCGGCCAAGATACAGAAGACGCGCT 2460
QY 3589 CACTTTGGGATCCGGATCTCGCTCTCTACTACAGCTTTCTTCAAGAGATTAAGGACAA 3648
D 2461 CACTTTGGGATCCGGATCTCGCTCTCTACTACAGCTTTCTTCAAGAGATTAAGGACAA 2520
QY 3649 TCGGATCGTCCGATCTGATTCAGACTCTGAAAGACGGTATCCGGGAAGTCGAGAAA 3708
D 2521 TCGGATCGTCCGATCTGATTCAGACTCTGAAAGACGGTATCCGGGAAGTCGAGAAA 2580
QY 3709 GAATATGCGAGGTTGTCAAAAACAAGAGATGAGACAGCAAGAACCCCTACCCCTGTC 3768
D 2581 GAATATGCGAGGTTGTCAAAAACAAGAGATGAGACAGCAAGAACCCCTACCCCTGTC 2640
QY 3769 CGGCTCAACCAAGTTTATGAAAAATGTCGCCCATCAAGCTTCAAGGCGATGACAAATCC 3828
D 2641 CGGCTCAACCAAGTTTATGAAAAATGTCGCCCATCAAGCTTCAAGGCGATGACAAATCC 2700
QY 3829 GGAGCAATATGATTTAAGGTGATCAGGTTGCTGAGGCTGTCCTCCGCGGACCGT 3888
D 2701 GGAGCAATATGATTTAAGGTGATCAGGTTGCTGAGGCTGTCCTCCGCGGACCGT 2760
QY 3889 GAGATGAATACATGGGCAATTGCTGAGGCTAGACAGCGCTTCAAGCTGTACTACCAAG 3948
D 2761 GAGATGAATACATGGGCAATTGCTGAGGCTAGACAGCGCTTCAAGCTGTACTACCAAG 2820
QY 3949 AGCCCCCAAGTTTGTGTGGCGAGATGCGCGGACAGACGTCCTGATCAATTAAGCGCAGATG 4008
D 2821 AGCCCCCAAGTTTGTGTGGCGAGATGCGCGGACAGACGTCCTGATCAATTAAGCGCAGATG 2880
QY 4009 ACAGACGACAGCCGGTGAAGGCGCCCGCGCTTGATGACCGGTTCAATGATGCGGCGTTG 4068
D 2881 ACAGACGACAGCCGGTGAAGGCGCCCGCGCTTGATGACCGGTTCAATGATGCGGCGTTG 2940
QY 4069 ATGCGCGATGAAGATTTAAGAACAGATATGTGGCCAGGCTGAGGGCGATGATCGAG 4128
D 2941 ATGCGCGATGAAGATTTAAGAACAGATATGTGGCCAGGCTGAGGGCGATGATCGAG 3000
QY 4129 TACCTTATCCGAGAGCTTATGAAGTCTGGCGATGATGATTTGATGAAATTTGTTTC 4188
D 3001 TACCTTATCCGAGAGCTTATGAAGTCTGGCGATGATGATTTGATGAAATTTGTTTC 3060
QY 4189 ACAGGAATGCGGATAT 4206
D 3061 ACAGGAATGCGGATAT 3078

```

## RESULT 4

```

AB254843
ID AB254843 standard; cDNA; 615 BP.
XX
AC AB254843;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide seq ID NO 3956.
XX
KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
XX expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002MO-IB000890.
XX
PR 30-MAR-2001; 2001JP-00098371.
XX

```

```

PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NOR-) NAT FOOD RES INST MIN AGRIC.
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
DR WPI; 2003-046817/04.
XX
PT Detection of expression of specific Aspergillus genes for monitoring the
PT fermentation and growth conditions of the fungus, using DNA probes.
XX
PS Claim 1; SEQ ID NO 3956; 48bp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (AB250888-AB254893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridizing
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 615 BP; 151 A; 148 C; 162 G; 154 T; 0 U; 0 Other;

```

```

Query Match 2.2%; Score 93.2; DB 8; Length 615;
Best Local Similarity 64.2%; Pred. No. 3.5e-17;
Matches 140; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```

```

QY 2869 GATGTCCTGTGGCGGATCCCGACCCCAATTCCTAGATATCAACGGGTGAGCA 2928
D 74 GACGTTCTTTACTTGGCTCCCGACATGTGCTGTGATATCCAGGACGAGAGCT 133
QY 2929 GTCTTAAGCCAGAGCTCCACAGTCTCAAGATGTATCATCTTCTTAAGGAGAT 2988
D 134 GTGTGAAGCAGAGCTCCGGGCAATTCAAAGATGATATGTTTCCCTACTGTGGAGCA 193
QY 2989 GTACCGCTTGCTTAAGAGCTATGTGTGAGACTAGACGGCGATGTGGCTGGCTTCG 3048
D 194 ACGCCACATCGCGACATGTATTCGGTGTGATTAAGACGGGATTAACAAGGGTTTC 253
QY 3049 TGGGATCCGAGATCGTCAATGTTTCGTCATGCGGA 3086
D 254 TGGGATCCAAATATTTGTCCAGAAATTTGTAATTCGA 291

```

## RESULT 5

```

ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001MO-IB001105.
XX
PR 22-JUN-2001; 2001MO-IB001105.
XX

```





CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

SQ Best Local Similarity 1.2%; Score 51.8; DB 8; Length 2000;

Query Match 1.2%; Score 51.8; DB 8; Length 2000;

Best Local Similarity 10.4%; Pred. No. 0.0004;

Matches 71; Conservative 309; Mismatches 301; Indels 4; Gaps 2;

QY 1454 TTGAGACCAAGTAGTGGCCGTTGTTCTCTGCTGTTAGTACATACCGGCAACT 1513  
 DB 711 WTKRAKSYRBRMYRMKMTYRMYRMSRMTBRMSRRKMGASMSKMMYRGA 652  
 QY 1514 CGCCTACTGCGCCCTTACTCTTGAAGCTGAAGCCGCTCATTTGAGAGGCTGTC 1573  
 DB 651 RSMYMSYKSCGAKCCCKTRMTSSYMSYKSMWSMTSMSTSM 592  
 QY 1574 GACTCACCCTGCGTGGTCTGATAGTGTTCGAGAT---CCTATACCGTGGCTTA 1610  
 DB 591 KGTBRKSMGMSMSRMYRMKMRKRYMRKMYRMTCTBRMCTRWGTYMTTSR 532  
 QY 1631 CGAGCACCAGCCCAAGTGTACCGGCTGTGACAGCAACACAGGTGCGGTGAGAG 1690  
 DB 531 MMYTGRYKARYTSRMYRMYRKYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 472  
 QY 1691 TCATCCAGTGGCTCAGCATGGGCAACATTTCTTTGAGGCGCCGATGCGGCTTCT 1750  
 DB 471 AYSMMYMYRKYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 412  
 QY 1751 TCAGCAAGATGCGGATACAGAACTCTCAGAGAGTTCGAGTCCGCGCGAGAGACC 1810  
 DB 411 YMSYKYSRCKYKMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 353  
 QY 1811 CGAAGCCATCATCAAGAGAGAGTCACTTCTTTGCGGAGACCGGATTAAGTTCGAC 1870  
 DB 352 KSYKKTCTWMCYKMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 293  
 QY 1871 CTGATGCTTCAACACAGATCTCTGCTCCGAGAGAGAACTCTAGAGAGAGACCG 1930  
 DB 292 MMRTAGMKRMSRMSRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 233  
 QY 1931 AGTTCAAGTGAATGCAATGCTGAGTGGCTCTGCAACTGCAACAAACACTTGAGCAGC 1990  
 DB 232 TSTRAKMRACRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 173  
 QY 1991 CGCACTCAAGTGTCTTCCCGTATCCAGTCTGCTGAGTGAAGATATATGATATGA 2050  
 DB 172 YGCAACMSGCMYMSYMSGCGCTGCMKMSKYSKCKKYSCTKYSYGYRCKMYK 113  
 QY 2051 CATTGAGGCTCAGCAGATGAGACACCAAGACCGATCTTCTTCACTTCAAGGACCTG 2110  
 DB 112 YSYKCYCYCYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 53  
 QY 2111 GCGAAGTGAATGATGAGGTGAG 2135  
 DB 52 SSGMYRMSKSKMYSKYSCKYTGS 28

RESULT 7  
 AA03159  
 ID AA03159 standard; DNA; 1164 BP.  
 XX  
 AC AA03159;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 02-FEB-1991 (first entry)  
 XX  
 DE Complete gene sequence of wildtype (WT) Streptomyces murinus glucose  
 DE isomerase (GI).  
 XX  
 KM Site-directed mutagenesis; glucose isomerase; high fructose corn syrup;  
 KM  
 XX  
 OS Streptomyces murinus.

XX Key Location/Qualifiers  
 FH CDS 1..1164  
 FT CDS /\*tag= a  
 FT  
 XX  
 XX  
 PN EP351029-A.  
 XX  
 PD 17-JAN-1990.  
 XX  
 XX 17-JUL-1989; 89EP-00201892.  
 XX  
 XX 15-JUL-1988; 88EP-00201539.  
 PR 04-NOV-1988; 88EP-00402789.  
 PR 17-JUL-1989; 89EP-00201893.  
 XX  
 PA (KONN) GIST-BROCADES NV.  
 PA (PLBZ) PLANT GENETIC SYSTEMS NV.  
 PA (GENV) GENENCOR INT INC.  
 XX  
 PI Luiten RGM, Quax WJ, Schuurhuiz FW, Mrabet N;  
 PI  
 XX  
 XX WPI; 1990-016368/03.  
 DR P-PSDB; AAR02215.  
 XX  
 PT New mutant glucose isomerase enzymes - obtd. by site-directed mutagenesis  
 PT of gene from Actinoplanes missouriensis, used for prodn. of high fructose  
 PT corn syrups.  
 PT  
 XX  
 PS Example; Fig 19; 56pp; English.  
 XX  
 CC The invention is a novel mutant glucose isomerase (GI) enzyme with  
 CC improved properties thru AA substitution. The GI is pref. derived from  
 CC Actinoplanes missouriensis. The substitn. is lys for Arg. or vice versa.  
 CC Specific lys or Arg residues are identified and site-directed  
 CC mutagenesis of the DNA sequence encoding the GI is performed. The novel  
 CC mutant gi shows 65 per cent or more sequence homology with the AA  
 CC sequence of WT A. missouriensis GI. Compared to wt enzyme it shows higher  
 CC conversion performance and improved thermostability and pH stability. The  
 CC example concerns the cloning and sequencing of GI from other bacterial  
 CC strains. Fig 21 displays the AA SSe to demonstrate the mutual homology.  
 CC (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 SQ Sequence 1164 BP; 178 A; 445 C; 382 G; 159 T; 0 U; 0 Other;

Query Match 1.2%; Score 51.6; DB 2; Length 1164;  
 Best Local Similarity 47.8%; Pred. No. 0.00033;  
 Matches 150; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 31 AGCCCGTGGAGAAATCATAAACGGCTCAATACGACTCAACCTGGGCTTCCAGTGT 90  
 DB 16 ACCCCGAGAGACAGTTCACCTTCGCTGTGTGAGACCGTGGCTGGAGGAGAGACCCG 75  
 QY 91 GTGCGAGACACACTCTACCCCGCCCAACCGCGGAAAGAGTGGCGAGTACGAGAGAT 150  
 DB 76 TTGCGGAGCGCACCGCCCGCCCTCGACCCGCTGAGAGCGGTGACGCGCTGGCGAG 135  
 QY 151 TTGCGTGCATGACAAGATCTACAGAGCCCTGAATCTTCTTACTGCGGAGAGATGAC 210  
 DB 136 CTGGGGCGCTTACGAGATGACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 195  
 QY 211 TCCCTGAACGAGCAGAGCAACCTTCTTCATGAGAGCCAAAGCTGCGAGCTGAACTGG 270  
 DB 196 ACCGAGCGGAGTGCAGACATCAAGCCCTTCCGCAAGCCCTGAGACCCAGCAGAGAG 255  
 QY 271 GTGCGCAAGAGCCAGAGCCGACCTGACAGGCTTCCGTGTCGCAAGAGAACCTCCCGGCG 330  
 DB 256 GTGCGCATGCGCACCAACCACTTACCCACACCCGCTTCAAGAGAGAGAGAGAGAGAG 315  
 QY 331 GCTACTGCGGAGCA 344  
 DB 316 GCCAAGACCGGGA 329



```
RESULT 8
AAS17845
ID AAS17845 standard; cDNA, 3731 BP.
XX
AC AAS17845;
XX
DT 08-MAY-2002 (first entry)
XX
DE Tomato RNA-directed RNA polymerase (RdRP) cDNA from clone HF.
XX
KW Tomato; RdRP; RNA-directed RNA polymerase; in vitro transcription;
KW cytosolic; virucide; RNA synthesis inhibitor; antibody; immunogen;
KW transgenic plant; transgenic animal; cancer; viral infection;
KW immunoprecipitation; immunolocalisation; RdRP/HF; mutant; gene therapy;
KW
XX
OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
FT mutation replace(31,T)
FT mutation /*tag= b
FT mutation replace(32,G)
FT CDS /*tag= c
FT /*tag= 194..3535
FT /*product= "RdRP protein"
FT /*EC_number= "2.7.7.48"
FT mutation replace(2017,G)
FT /*tag= d
FT mutation replace(2264,T)
FT /*tag= e
FT mutation replace(3597,A)
FT /*tag= f
FT mutation replace(3599,A)
FT /*tag= g
XX
PN US2001023067-A1.
XX
PD 20-SEP-2001.
XX
PF 08-FEB-2001; 2001US-00782874.
XX
PR 05-MAR-1997; 97US-00811583.
XX
PA (WASS/) WASSENEGGER M.
PA (RIED/) RIEDEL L.
PI Wassenegger M, Riedel L, Schiebel W, Sanger HL;
XX
DR WPI: 2001-595798/67.
XX
PT New nucleic acid molecule encoding a polypeptide having the enzymatic
PT activity of RNA-directed RNA polymerase, for modulating gene expression
PT and treating cancer and virus infection in human and animals.
XX
PS Example 3; Page; 34pp; English.
XX
CC This sequence represents a cDNA encoding the tomato RNA-directed RNA
CC polymerase (RdRP/HF) of the invention. The invention comprises the
CC nucleic acid and protein sequences of RdRP. The protein of the invention
CC can catalyse in vitro transcription of short single stranded RNAs into
CC DNA molecules, this transcription can be either primed by RNA or DNA
CC oligonucleotides or be unprimed. The protein may have cytosolic or
CC virucide activities. The sequences of the invention may be used in gene
CC therapy or as an RNA directed RNA synthesis inhibitor. The RdRP cDNA
CC sequence and a template nucleic acid molecule derived from a nucleic acid
CC molecule which causes a disease are useful for treating a disease caused
CC by the undesired expression or overexpression of a nucleic acid molecule
CC in a human, rat or mouse, by administering the molecules. This system can
CC be used in the preparation of a pharmaceutical composition and for
CC inhibiting expression of any desired gene by transferring the RdRP system
CC to organisms that either lack a comparable mechanism or do not
```

```
CC sufficiently express their own RdRP. An antibody or an antagonist or
CC inhibitor to the protein are useful for inhibiting RNA directed RNA
CC synthesis and for ensuring stable heterologous, gene expression in
CC transgenic organisms. The sequence is useful for probes and/or for the
CC control of gene expression, as primers for amplification of nucleic acid
CC molecules and as tools for the detection of expression of the cDNA
CC molecules. Additionally, nucleotide and protein sequences are useful for
CC suppression of undesired gene expression in humans and animals. The RdRP
CC is useful as a therapeutic agent for the control of cancer and virus
CC infection in humans and animals and the antibody is useful for
CC immunoprecipitation or immunolocalisation of the protein, identification
CC of polypeptides interacting with it and screening expression libraries.
CC Note: This sequence is not shown in the specification but was created by
CC the indexer from the wild type RdRP sequence shown in AAS17837 and the
CC information shown in table 1
XX
SQ Sequence 3731 BP; 1062 A; 672 C; 848 G; 1149 T; 0 U; 0 Other;
XX
Query Match 1.2%; Score 49; DB 4; Length 3731;
Best Local Similarity 51.6%; Pred. No. 0.0043;
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
OY 2850 TACACTCCTATGCGACTGTGATGTCCTGTCGGCCGATGCCACCCATTCCCTAGTGA 2909
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2425 TAATTTCATTCTGAAGGGAATGCTGTGCAAAAATCCATGCTGCATCCTGGTGA 2484
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2910 TATCCACGGGTTGAGACAGCTTCAAGCCAGACCTCCACAGTCTCAAGATGTAATCAT 2969
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2485 TATTCGTGTTTAAAGCTGTAATGTTTCAGACCTCCACACATGAGTATGTTGTGT 2544
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2970 CTTCTCTACTAAGAGATGATACCGCTTGCTAAGACCTATCTGTGAGACTACGACGG 3029
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2545 ATTCCCTCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2604
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3030 CGATATGAGCTGAGCTGCTGCTGGGATCCGAGATGCTC 3066
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2605 GGATATCTACTTGTGTTGCTGGGATCAAGACATGATC 2641
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
RESULT 9
AAS17837
ID AAS17837 standard; cDNA, 3731 BP.
XX
AC AAS17837;
XX
DT 08-MAY-2002 (first entry)
XX
DE Tomato RNA-directed RNA polymerase (RdRP) cDNA.
XX
KW Tomato; RdRP; RNA-directed RNA polymerase; in vitro transcription;
KW cytosolic; virucide; RNA synthesis inhibitor; antibody; immunogen;
KW transgenic plant; transgenic animal; cancer; viral infection;
KW immunoprecipitation; immunolocalisation; ss; gene therapy.
XX
OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
FT CDS /*tag= a
FT /*tag= 194..3538
FT /*product= "RdRP protein"
FT /*EC_number= "2.7.7.48"
XX
PN US2001023067-A1.
XX
PD 20-SEP-2001.
XX
PF 08-FEB-2001; 2001US-00782874.
XX
PR 05-MAR-1997; 97US-00811583.
XX
PA (WASS/) WASSENEGGER M.
PA (RIED/) RIEDEL L.
XX
```

PI Masenegger M, Riedel L, Schiebel W, Sanger HL;  
XX WPI; 2001-595798/67.  
DR P-PSDB; AAU10006.  
XX  
XX New nucleic acid molecule encoding a polypeptide having the enzymatic  
PT activity of RNA-directed RNA polymerase, for modulating gene expression  
PT and treating cancer and virus infection in human and animals.  
PS Claim 1; Page 15-19; 34pp; English.  
XX  
XX This sequence represents a cDNA encoding the tomato RNA-directed RNA  
CC polymerase (RdRp) protein of the invention. The invention comprises the  
CC nucleic acid and protein sequences of RdRp. The protein of the invention  
CC can catalyze in vitro transcription of short single stranded RNAs into  
CC DNA molecules, this transcription can be either primed by RNA or DNA  
CC oligonucleotides or be unprimed. The protein may have cytosolic or  
CC virucide activities. The sequences of the invention may be used in gene  
CC therapy or as an RNA directed RNA synthesis inhibitor. The RdRp cDNA  
CC sequence and a template nucleic acid molecule derived from a nucleic acid  
CC molecule which causes a disease are useful for treating a disease caused  
CC by the undesired expression or overexpression of a nucleic acid molecule  
CC in a human, rat or mouse, by administering the molecules. This system can  
CC be used in the preparation of a pharmaceutical composition and for  
CC inhibiting expression of any desired gene by transferring the RdRp system  
CC to organisms that either lack a comparable mechanism or do not  
CC sufficiently express their own RdRp. An antibody or an antagonist or  
CC inhibitor to the protein are useful for inhibiting RNA directed RNA  
CC synthesis and for ensuring stable heterologous gene expression in  
CC transgenic organisms. The sequence is useful for probes and/or for the  
CC control of gene expression, as primers for amplification of nucleic acid  
CC molecules and as tools for the detection of expression of the cDNA  
CC molecules. Additionally, nucleotide and protein sequences are useful for  
CC suppression of undesired gene expression in humans and animals. The RdRp  
CC is useful as a therapeutic agent for the control of cancer and virus  
CC infection in humans and animals and the antibody is useful for  
CC immunoprecipitation or immunolocalisation of the protein, identification  
CC of polypeptides interacting with it and screening expression libraries  
XX  
SQ Sequence 3731 BP; 1064 A; 669 C; 849 G; 1149 T; 0 U; 0 Other;  
Query Match 1.2%; Score 49; DB 4; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.0043;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGTGATGCTCTGTCGCGCGATCCCGACCCATTTCCTAGTGA 2909  
Db 2425 TAATTTTCATCTGAAAGGAAATGTTGTCAAAAAATTCATGCTTGATCTGCTGTA 2484  
QY 2910 TATCCAAAGGGTTCGACGACTTTCAGCCAGAGCTCCAGATGTCAGAGATGTAATCAT 2969  
Db 2485 TATTCGTGTTTAAAGGCTGTAATGTTGACGCGCTGCACACATGTAGATTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATACCGCTTGCTAAGAGCTATCTGGTGGAGACTACGACGG 3029  
Db 2545 ATTCCCTCAGAAAAGAAAAGACCTCATCCGAATGAATGTTCTGGAGATGATTTGGATGG 2604  
QY 3030 CGATATGCGCTGGGTCTGCTGGGATCCGAGATCGTC 3066  
Db 2605 GGATATCTACTTTGTTGCTGGGATCAAGACATGATC 2641  
RESULT 10  
AAD04370 ID AAD04370 standard; cDNA; 3731 BP.  
XX  
XX AAD04370;  
XX  
XX 04-JUL-2001 (first entry)  
XX  
XX Tomato RNA-directed RNA polymerase (RdRp) cDNA.  
XX  
XX  
XX Tomato; gene therapy; RNA-directed RNA polymerase; RdRp; gene expression;  
KM

KW transgenic plant; tissue culture; plant breeding; therapy; C-protein; ss.  
XX  
XX Lycopersicon esculentum.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 194..3538  
FT /tag= a  
FT /product= "Tomato C-protein having RNA-directed RNA  
FT polymerase (RdRp) activity"  
PS US6218142-B1.  
XX  
XX 17-APR-2001.  
PD  
XX  
XX 05-MAR-1997; 97US-00811583.  
PF  
XX  
XX 05-MAR-1997; 97US-00811583.  
PR  
XX  
XX (WASS/) MASSENEGGER M.  
PA (RIED/) RIEDEL L.  
PI Masenegger M, Riedel L, Schiebel W, Sanger HL;  
XX WPI; 2001-289830/30.  
DR P-PSDB; AAE00897.  
XX  
XX New nucleic acid molecules encoding polypeptides with RNA-directed RNA  
PT polymerase enzymatic activity, useful in modulating gene expression in  
PT plants, humans and animals, as well as in plant cell/tissue cultures or  
PT plant breeding.  
XX  
PS Claim 1; Col 25-34; 31pp; English.  
XX  
XX The present sequence is a cDNA encoding tomato C-protein having RNA-  
CC directed RNA polymerase (RdRp) activity. This protein is capable of RNA-  
CC directed RNA synthesis, thus using RNA as a template for synthesising  
CC complementary RNA molecules. RdRp nucleic acid is useful for modulating  
CC gene expression in plants, humans and animals. This may lead to various  
CC physiological, developmental and/or morphological changes. Transgenic  
CC plants containing RdRp nucleic acid is especially useful in plant cell or  
CC tissue cultures and in plant breeding. RdRp is useful in gene therapy,  
CC particularly for treating a disease that is caused by the undesirable  
CC expression or overexpression of a gene  
XX  
SQ Sequence 3731 BP; 1064 A; 669 C; 849 G; 1149 T; 0 U; 0 Other;  
Query Match 1.2%; Score 49; DB 4; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.0043;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGTGATGCTCTGTCGCGCGATCCCGACCCATTTCCTAGTGA 2909  
Db 2425 TAATTTTCATCTGAAAGGAAATGTTGTCAAAAAATTCATGCTTGATCTGCTGTA 2484  
QY 2910 TATCCAAAGGGTTCGACGACTTTCAGCCAGAGCTCCAGATGTCAGAGATGTAATCAT 2969  
Db 2485 TATTCGTGTTTAAAGGCTGTAATGTTGACGCGCTGCACACATGTAGATTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATACCGCTTGCTAAGAGCTATCTGGTGGAGACTACGACGG 3029  
Db 2545 ATTCCCTCAGAAAAGAAAAGACCTCATCCGAATGAATGTTCTGGAGATGATTTGGATGG 2604  
QY 3030 CGATATGCGCTGGGTCTGCTGGGATCCGAGATCGTC 3066  
Db 2605 GGATATCTACTTTGTTGCTGGGATCAAGACATGATC 2641  
RESULT 11  
ACH16195 ID ACH16195 standard; cDNA; 495 BP.  
XX  
XX ACH16195;  
XX

DT 13-OCT-2003 (first entry)  
XX  
DE Human adult heart cDNA #509.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2003073623-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 30-JUL-2001; 2001US-00918995.  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI WPI; 2003-615964/58.  
DR  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
PS Claim 1; SEQ ID NO 3407; 44pp; English.  
XX  
CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 495 BP; 137 A; 123 C; 123 G; 111 T; 0 U; 1 Other;  
Query Match 1.1%; Score 45.6; DB 9; Length 495;  
Best Local Similarity 51.5%; Pred. No. 0.013; 99; Indels 0; Gaps 0;  
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 2506 AATGCCCGCATGTAATTCGCAATGGCTTACGAGATTATTCAGTCGCGCAACTGCA 2565  
DB 287 AATACCAAGTGGCTTTCTTCTACTGACATGATGACATATCAATGCCCAACAAA 346  
QY 2566 GTACAGCCAGCGCGTGTCTTTTCTGCGCTACTGACATGACAGAGACACTG 2625  
DB 347 GACCTCTTGACCCCTTGCACTTGCCTGCGAAAGACAGCAAGCAATACCTTAATA 406  
QY 2626 AACTCTGATGAACAGTGGCTTCATCCCAAGAAAGCAAAAGTACTTCAGACATCGCC 2685  
DB 407 CTCCTCCGATGAACCCCTTACGTCAACAGGCGCTGAAAAACAACTTGGAAAGAACTCA 466  
QY 2686 TGGGATCTTCAAAACGGGAAATGT 2709  
DB 467 TTGTATTTGCCAGAGACAAGT 490

RÉSULT 12  
AAC84447  
ID AAC84447 standard; cDNA; 2710 BP.  
XX  
AC AAC84447;  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of mdt cDNA clone ID No: 237703.2.  
XX  
KW Molecule for disease detection and treatment; mdt; human; cirrhosis;  
KW cell proliferative disorder; arteriosclerosis; bursitis; hepatitis;  
KW connective tissue disease; myelofibrosis; cancer; leukemia; autoimmune;  
KW acquired immunodeficiency syndrome; AIDS; Addison's disease; allergy;  
KW anemia; bronchitis; gout; Hashimoto's thyroiditis; multiple sclerosis;  
KW cytostatic; immunomodulator; anti-inflammatory; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200075298-A2.  
XX  
PD 14-DEC-2000.  
XX  
PF 01-JUN-2000; 2000WO-US015344.  
XX  
PR 03-JUN-1999; 99US-0137412P.  
PR 05-AUG-1999; 99US-0147500P.  
PR 05-AUG-1999; 99US-0147501P.  
PR 05-AUG-1999; 99US-0147542P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;  
PI Bratcher SR, Dufour GE, Cohen HJ, Rosen BH, Chalup MS, Hillman JV;  
PI Jones AL, Yu JV, Greenwalt LB, Panzer SR, Roseberry AM, Wright RJ;  
PI Daniels SE;  
XX  
DR WPI; 2001-071068/08.  
XX  
PT New polypeptide and polynucleotide molecules for disease detection and  
PT treatment are useful in diagnosis and gene therapy of proliferative  
PT disorders (e.g. breast cancer), autoimmune and inflammatory disorders  
PT (e.g. AIDS or allergy).  
XX  
PS Claim 1; Page 90; 99pp; English.  
XX  
SQ Sequences AAC84445-C84458 are new isolated polynucleotide molecules for  
CC disease detection and treatment (mdt). The mdt polynucleotides are  
CC useful for diagnosing or treating disorders associated with disease  
CC detection and treatment molecules. These diseases include cell  
CC proliferative disorders (e.g. arteriosclerosis, bursitis, cirrhosis,  
CC hepatitis, mixed connective tissue disease, myelofibrosis, or cancers  
CC such as leukemia, or breast or brain cancers), autoimmune or inflammatory  
CC disorders (e.g. acquired immunodeficiency syndrome (AIDS), Addison's  
CC disease, allergies, anemia, autoimmune haemolytic anemia, bronchitis,  
CC gout, Hashimoto's thyroiditis, multiple sclerosis or cancers). The mdt  
CC are especially useful for somatic or germline gene therapy. The mdt may  
CC also be used to detect the presence of, or to quantify the amount of, an  
CC mdt-related polynucleotide in a sample. The mdt are also useful for  
CC isolating full length cDNA sequences utilizing hybridization and/or  
CC amplification procedures, for generating hybridization probes useful in  
CC chromosomal mapping of naturally occurring genomic sequences, as  
CC molecular weight markers, or for monitoring the progress of disorders  
CC associated with abnormal levels of mdt expression or evaluating the  
CC efficacy of a particular treatment. The recombinant nucleic acids are  
CC useful as part of a viral vector (e.g. based on a vaccinia virus) for  
CC vaccinating a mammal and inducing a protective immunological response in  
CC the mammal. The MDT polypeptides are useful for screening molecules that  
CC bind to or are bound by the polypeptides, i.e. for screening agonists,  
XX antagonists or modulators of MDT

Sequence 2710 BP; 856 A; 516 C; 450 G; 880 T; 0 U; 8 Other;  
Query Match 1.1%; Score 45.6; DB 4; Length 2710;  
Best Local Similarity 51.5%; Pred. No. 0.039;  
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 2506 AATGGCCAGTGGAAATTTGCGCAATGGTTACAGAGTTATTCAGTCCGCACTCGA 2565  
DB 1010 AATCCAGAGTGGCTTTCTTTCTACTCAGCATATGCAATATCAATGCCCCAACA 1069  
QY 2566 GTCAGCCAGGCGCTGTGCTTTCTTCTGTGGCTTACCTGACAGTCAAGAGACACTG 2625  
DB 1070 GGCCTCTTGACCCCTTGTCATCTTGCTGTGGAAACAGACAGCAAGATACCTTGAA 1129  
QY 2626 AACTTCTTGATGAACAGTGGTTCATCCCAAGAGAAAAGTATCTTGCAAGACATGCC 2685  
DB 1130 CTCCTCTGATGAACCGTTAGTCAAAACAGGCGCTGAAAAACAATTGGAAGAACTGCA 1189  
QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
DB 1190 TTGTATATTGCCAGAGACAAAGT 1213  
RESULT 13  
ID ABL71773 standard; cDNA; 282 BP.  
XX ABL71773;  
AC ABL71773;  
XX 14-MAY-2002 (first entry)  
DE Corn tassael-derived polynucleotide (cdps) SEQ ID NO:1147.  
XX  
XX Corn; corn tassael-derived polynucleotide; cdps; hybrid breeding; CDPS;  
KW inheritance; characteristic; growth; development; disease resistance;  
KW environmental adaptability; quality; yield; molecular marker;  
KW multigene trait; plant breeding; corn tassael; gene; ss.  
XX  
XX Zea mays.  
OS  
XX US2001051335-A1.  
PN 13-DEC-2001.  
PD  
XX 16-APR-1999; 99US-00294093.  
PF  
XX 21-APR-1998; 98US-0082567P.  
PR  
XX (LALG/) LALGUDI R V.  
PA (ITOL/) ITO L Y.  
PA (SHER/) SHERMAN B K.  
XX  
XX Lalgudi RV, Ito LY, Sherman BK;  
PI WPI; 2002-163647/21.  
DR  
XX  
XX Novel purified corn tassael-derived polynucleotide useful for determining  
PT altered gene expression, to recover regulatory elements and to follow  
PT inheritance of desirable characteristics through hybrid breeding  
PT programs.  
XX  
XX Claim 1; SEQ ID NO 1147; 201pp; English.  
PS  
XX The present sequence describes a purified corn tassael-derived  
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence  
CC selected from those given in ABL70627 to ABL76833. The cdps sequences  
CC encode corn tassael-derived polypeptides (CDPS). The cdps sequences (1)  
CC can be used for determining altered gene expression, to recover  
CC regulatory elements and to follow inheritance of desirable  
CC characteristics through hybrid breeding programs. (1) are also useful in  
CC the evaluation, and alteration of desired characteristics associated with  
CC growth and development, disease resistance, environmental adaptability,  
CC quality and yield, and as molecular markers for studying inheritance of

CC multigene traits in a plant breeding program. (1) can be used to produce  
CC a tassael-specific profile of gene transcription, a transcript image, to  
CC clone regulatory elements for use in transformation vectors, to express a  
CC polypeptide, to identify, isolate or extend identical or related corn  
CC tassael nucleic acid sequences from DNA libraries, in nucleic acid  
CC hybridisation or amplification technologies, as query sequences to  
CC determine homology of known sequences, as probe for use in Southern or  
CC Northern hybridisation, and to identify the presence of and/or to  
CC determine the degree of similarity between two (or more) nucleic acid  
CC sequences  
XX  
XX Sequence 282 BP; 78 A; 67 C; 70 G; 67 T; 0 U; 0 Other;  
Query Match 1.0%; Score 44; DB 6; Length 282;  
Best Local Similarity 52.1%; Pred. No. 0.029;  
Matches 98; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 2867 GTGATGTCCTCGTGGCGCATCCCGAGCCATTTCCTAGTATATTCACAGGGTTCGAG 2926  
DB 16 GTACTATCGTAAAGCAAGAAATCCATGCTTCATCCAGGAGTGTCCGATCTTGAAAG 75  
QY 2927 CAGTCTTCAAGCCAGAGCTCCACAGTCTCAAGGATGTAATCATCTTCTACTAAAGAG 2986  
DB 76 CTGTGATGTCGCTGAACTGCATCACCCTTGTATGCTTGTGCTTCCCAAGAAAGTG 135  
QY 2987 ATGTACCGCTTGTGAAGAGTATCTGTGTGAGACTACAGCGGCGATATGCGGTCT 3046  
DB 136 AGAGCCGCGACGCAATGAAGCATCTGGAGTGATCTTGATGGGATATATCTTGTA 195  
QY 3047 GCTGGAT 3054  
DB 196 CATGGAT 203  
RESULT 14  
ID AAS33132 standard; cDNA; 1050 BP.  
XX AAS33132;  
AC AAS33132;  
XX 04-DEC-2001 (first entry)  
DE DNA encoding human secreted protein, Seq ID No 91.  
XX  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200155326-A2.  
PN 02-AUG-2001.  
PD  
XX 17-JAN-2001; 2001MO-US001347.  
PF  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-018464P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220966P.  
PR 26-JUL-2000; 2000US-0220966P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231966P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234998P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249267P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-451931/48.  
P-PSDB; AAU20423.  
  
New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.  
  
XX PS Claim 1; SEQ ID NO 91; 753bp; English.  
XX The invention relates to novel isolated nucleic acid molecules (I)  
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
CC the prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression of secreted proteins. (I) and complementary  
CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
CC of similar nucleic acid sequences in samples, and so which patients may  
CC be in need of restorative therapy. (II) may also be used as antigens in

CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of the secreted  
CC proteins. The anti-(II) antibodies and antagonists may also be used to  
CC down regulate expression and activity of (II). The anti-(II) antibodies  
CC may also be used as diagnostic agents for detecting the presence of (II)  
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and  
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular  
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists  
CC and antibodies can also be used to promote wound healing, maintain organs  
CC before transplantation, and support cell culture of primary tissues.

Query Match 1.0%; Score 44; DB 4; Length 1050;

Best Local Similarity 51.0%; Pred. No. 0.067;  
Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCAGTGGAAATTTCCCATATGGCTTACGAGATTATTCATCCGCGCAATCGA 2565  
DB 570 AATACCGAGAGTGGCTTTCTTTCTTACTGACATGATGCAATATCAATGCCAATAA 629  
QY 2566 GTACGACGCGCGGTGTGCTTTCTTCTGCTGGCTACCTGACATCAAGAGACACTG 2625  
DB 630 GGCTCTTGACCCCTTGATCTTGCTGCTGGGACAGACAGACAGATACCTTGAA 689  
QY 2626 AACTCTTGATGAACAGTGGGTTGATCCCAAGAAAGAAAGTATCTGCAAGCATGCC 2685  
DB 690 CTCCTCTGATGAACCGTTACGTCAACCAAGGCTGAATAAACAATTGGAAGAACTGCA 749  
QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
DB 750 TTGATATTGCCAGAGACAAGT 773

## RESULT 15

AAF62769  
ID AAF62769 standard; cDNA: 1318 BP.

XX AAF62769;

DT 03-MAY-2001 (first entry)

XX Human FGIF cDNA.

XX Human; FGIF; foetal globin inducing factor; antihaemic; antianginal;  
KW antisickling; analgesic; sickle cell anaemia; thalassaemia;  
KW haemoglobinopathy; ss.

XX Homo sapiens.

XX US6184343-B1.

PN 06-FEB-2001.

XX 05-DEC-1997; 97US-00986304.

XX 06-DEC-1996; 96US-0033247P.

XX (STAM/) STAMATOYANNOPOLIOS G.

PA (PAPA/) PAPAYANNOPOULOU T.

XX Stamatoyanopoulos G, Papayannopoulou T;

XX WPI: 2001-225816/23.

DR P-PDB; AAB71861.

XX Novel fetal globin inducing factor capable of regulating expression of  
PT globin gene, for treating globin disorders such as sickle cell anemia,

PT thalassemia and for screening modulators of fetal globin gene expression.  
XX  
XX Example; Fig 1; 16pp; English.

XX The present sequence encodes human foetal globin inducing factor (FGIF).  
CC FGIF is useful for regulating the expression of globin genes such as  
CC foetal globin. It may be used in the prophylactic or therapeutic  
CC treatment of globin-associated disorders such as sickle cell anaemia and  
CC thalassaemia. Symptoms associated with globin disorders treatable by FGIF  
CC include anaemia, tissue hypoxia, organ dysfunction, abnormal haematocrit  
CC values, ineffective erythropoiesis, abnormal reticulocyte count, abnormal  
CC iron load, splenomegaly, hepatomegaly, impaired peripheral blood flow,  
CC dyspnea, increased haemolysis, jaundice, anaemic crises and pain such as  
CC angina pectoris. FGIF is also useful for screening for agents that  
CC modulate its function. Agents that enhance FGIF function can be used to  
CC treat haemoglobinopathies and beta-thalassaemia syndromes. A relatively  
CC small number of FGIF molecules transactivates expression of endogenous  
CC foetal globin genes. This method is highly advantageous over prior art  
CC methods, using vectors carrying either human foetal or human adult globin  
CC genes for expression of therapeutic polypeptides, which require a  
CC consistent production of high levels of globin mRNA in the target cells  
XX

SQ Sequence 1318 BP; 414 A; 245 C; 264 G; 395 T; 0 U; 0 Other;

Query Match 1.0%; Score 44; DB 5; Length 1318;

Best Local Similarity 51.0%; Pred. No. 0.077;  
Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCAGTGGAAATTTCCCATATGGCTTACGAGATTATTCATCCGCGCAATCGA 2565  
DB 592 AATACCGAGAGTGGCTTTCTTTCTTACTGACATGATGCAATATCAATGCCAATAA 651  
QY 2566 GTACGACGCGCGGTGTGCTTTCTTCTGCTGGCTACCTGACATCAAGAGACACTG 2625  
DB 652 GGCTCTTGACCCCTTGATCTTGCTGCTGGGACAGACAGACAGATACCTTGAA 711  
QY 2626 AACTCTTGATGAACAGTGGGTTGATCCCAAGAAAGAAAGTATCTGCAAGCATGCC 2685  
DB 712 CTCCTCTGATGAACCGTTACGTCAACCAAGGCTGAATAAACAATTGGAAGAACTGCA 771  
QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
DB 772 TTGATATTGCCAGAGACAAGT 795

Search completed: August 10, 2005, 19:57:15  
Job time : 2065 secs



NAME/KEY: CDS  
LOCATION: 1..1164  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /EC number=5.3.1.5  
OTHER INFORMATION: /product="xylose isomerase (glucose isomerase)"  
OTHER INFORMATION: /evidence=EXPERIMENTAL  
OTHER INFORMATION: /standard\_name="D-xylose ketol isomerase"  
US-07-640-476-6

Query Match 1.2%; Score 51.6; DB 1; Length 1164;  
Best Local Similarity 47.8%; Pred. No. 2.5e-05;  
Matches 150; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 31 AGCCCGCTCGAGAAATCATTAACCGGCTCATATACGACTACAACTGGGCTTCCAGTGT 90  
DB 16 ACCCCGAGAGACAGATTACCTTGGTGTGAGACCTGGCTGGCAGGAGGAGACCCG 75  
QY 91 GTGCGACACAACTCTACACCCCGACCGCGGAGAGAGCTGGCGGAGAGAGAGAT 150  
DB 76 TTGCGGACGCGACCCCGCCCGCTGAGACCGGTGAGAGCGGTGAGCGCTGGCGAG 135  
QY 151 TTGCGTCCGATGACAAATCTACAGAGCCCTGAATCTTCTACTGCGGAGAGATGAC 210  
DB 136 CTGGGGCCCTACAGAGAGACCTTCCAGAGAGAGACCTGATCCCTTGGGTCTCCGAC 195  
QY 211 TCCTGACACGAGGAGAGCACTTTCTTATGAGGCGCAAGCTGCGAGCTGCACTGG 270  
DB 196 ACCGAGGCGAGTGCATACATACGCTTCGCGGAGCGCTGAGCGCACCGGATGACG 255  
QY 271 GTGCGCAAGCGGCGGCGGAGCCCTGACAGCGCTTCGCTGTCAGAGAACTCCCGCGGC 330  
DB 256 GTGCGCATGGCCACACCAACTTTTACCCACCCGCTTTCAGAGAGCGGCGCTTACC 315  
QY 331 GCTACTCCGCGCA 344  
DB 316 GCCAAGACCGCGCA 329

RESULT 2  
US-08-811-583-1  
Sequence 1, Application US/08811583  
Patent No. 6218142

GENERAL INFORMATION:  
APPLICANT: Maassenegger, Michael  
APPLICANT: Riedel, Leonhard  
APPLICANT: Schiebel, Winfried  
APPLICANT: Sanger, Heinz  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
TITLE OF INVENTION: POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN  
TITLE OF INVENTION: RNA-DIRECTED RNA POLYMERASE (RDRP)  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,583  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: MPG-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3731 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Tomato  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 194..3535  
US-08-811-583-1

Query Match 1.2%; Score 49; DB 3; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.00041;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2850 TACACTCTATCGAGCTGATGTCTCTGTGGCGGATCCCAAGCCATTTCCTAGTGA 2909  
DB 2425 TAATTTCATTCTGAAGGAAATGTGTTGGCAAAAATCCATGCTTGCATCTGCTGA 2484  
QY 2910 TATCCAAAGGCTTCAGCAAGCTTCAAGCAGCTCCACAGTCTCAAGATGTAATCAT 2969  
DB 2485 TATTCGTGTTTAAAGCTGTAAATGTTGAGGCGCTGCACACATGATGATGTTGT 2544  
QY 2970 CTTCCTACTAAGAGAGATGATACCGCTTGCTAAGAACTATCTGTGAGACTACGACG 3029  
DB 2545 ATTCCCTCAAAAAGAAAAAGACCTCATCGAATGATGTTCTGGAGATGATTTGATGG 2604  
QY 3030 CGATATGCGCTGGGTCTGCTGGGATCCGAGATGCTC 3066  
DB 2605 GGATATCTACTTGTGTTGCTGGGATCAAGACATGATC 2641

RESULT 3  
US-08-986-304-1  
Sequence 1, Application US/08986304A  
Patent No. 6184343

GENERAL INFORMATION:  
APPLICANT: Stamatoyanopoulos, George  
APPLICANT: Papayannopoulos, Thalia  
APPLICANT: Yang, Yi  
TITLE OF INVENTION: FETAL GLOBIN INDUCING FACTOR  
FILE REFERENCE: 64657.0102  
CURRENT APPLICATION NUMBER: US/08/986,304A  
CURRENT FILING DATE: 1997-12-05  
EARLIER APPLICATION NUMBER: 60/033,247  
EARLIER FILING DATE: 1996-12-06  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1318  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-986-304-1

Query Match 1.0%; Score 44; DB 3; Length 1318;  
Best Local Similarity 51.0%; Pred. No. 0.0078;  
Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCCAAGTGAATTTGCGCAATGGTTTACGAGATTAATTCAGTCCGCACTCGA 2565  
DB 592 AATACCAAGAGTGTCTTCTTACTGACGATGATGACATATCAATGCCCAACAAA 651  
QY 2566 GTAGACGAGCGCGGTGCTTTTCTTCTGGGCTACCTGACAGTCAAGAGAGACACTG 2625  
DB 652 GGCCTCTTGACCCCTTGCACTTGTGCTGGGAGACAGACGAAAGATACCTAGAA 711  
QY 2626 AACTCTTGATGAAGAGTGGGTTCGATCCCAAGCAAAAGTACTGACATCGCC 2685  
DB 712 CTCCTCTATGAAACCTTATGCTAAACCGGCTGAAAAACAATTTGGAGAACTGCA 771





```

RESULT 7
US-09-103-840A-1
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ. ID NOS.: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37RV
US-09-103-840A-1

```

	Query Match	Best local Similarity	1.0%;	Score 40.8;	DB 3;	Length 4411529;
	Matches 105;	Conservative	49.5%;	Pred. No. 37;	Mismatches 107;	Indels 0; Gaps 0;
QY	3952	CCCAAGTTCGTGTGGCAGATGGCGGGCAGACAGCTGCAGTACATTAAAGCGCAGATAGACG				4011
Db	426675	CCCGAGTTGGTCCGCCGCGATGCCACCTTGCTGTGCGCGGATAGGCCACGCCGAAACCGG				426734
QY	4012	AGCAGACCCGGGTAAAGGGCCCCGGCGTTGATAGACCGGGTTCATGTATGCGGGCTTATG				4071
Db	426735	CCGGTCCCATGTTGCCCATGCCGAGTTGCCGGTCCCGAGTTGAACAAACCCGGTGTG				426794
QY	4072	CCGGAATAAGATTTTACGAGCAGTATGTGCGCAGGCTTGAGAGGGCAGTATCCGAGTAC				4131
Db	426795	CCGGTCCCGGAGTTGAACAAGCCGGGTGTGGCGGTGCCGAGTTGAAGAAACCCGTTGTG				426855
QY	4132	CTGTATCCCGGAGGCTCTATGAATGCTCGGGCCGA				4163
Db	426855	CCGGCCGCCGAGTTCCAGGAGCTGGAAGCCGGA				426886

```

RESULT 8
US-09-270-767-12025/C
; Sequence 12025, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7336-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12025
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12025

```

Query Match	1.0%	Score 40.2	DB 4	Length 803
Best Local Similarity	70.1%	Pred. No. 0.091		
Matches	54	Conservative	0	Mismatches 23
				Indels 0
				Gaps 0
QY	4109	TGAGGGCGATGCATCGGAGTACCTCTGATCCGAGGTCTATGAATGCTGGGCGCATGATG	4168	
Db	96	TGGAGGACCATGTCGCGGTGGCCCATGTCGGGGAGGTCACATGTCGTGTGGGCCATGATG	37	
QY	4169	ATTTTGATGCATTCGT	4185	

Db 36 ATGGTGGTGGTCTGGT 20

```

1 RESULT 9
2 US-09-107-532A-1706
3 ; Sequence 1706, Application US/09107532A
4 ; Patent No. 6583275
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
8 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
9 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
10 ;
11 ; NUMBER OF SEQUENCES: 7310
12 ;
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
15 ; STREET: 100 Beaver Street
16 ; City: Waltham
17 ; STATE: Massachusetts
18 ; COUNTRY: USA
19 ; ZIP: 02354
20 ;
21 ; COMPUTER READABLE FORM:
22 ; MEDIUM TYPE: CD/ROM ISO9660
23 ;
24 ; COMPUTER: PC
25 ; OPERATING SYSTEM: <Unknown>
26 ; SOFTWARE: ASCII
27 ;
28 ; CURRENT APPLICATION DATA:
29 ; APPLICATION NUMBER: US/09/107,532A
30 ; FILING DATE: 30-Jun-1998
31 ;
32 ; PRIOR APPLICATION DATA:
33 ; APPLICATION NUMBER: 60/085,598
34 ; FILING DATE: 14 May 1998
35 ; APPLICATION NUMBER: 60/051571
36 ; FILING DATE: July 2, 1997
37 ;
38 ; ATTORNEY/AGENT INFORMATION:
39 ; NAME: Arinello, Pamela Deneke
40 ; REGISTRATION NUMBER: 40,489
41 ; REFERENCE/DOCKET NUMBER: GTC-012
42 ;
43 ; TELECOMMUNICATION INFORMATION:
44 ; TELEPHONE: (781) 893-5007
45 ; TELEFAX: (781) 893-8277
46 ;
47 ; INFORMATION FOR SEQ ID NO: 1706:
48 ; SEQUENCE CHARACTERISTICS:
49 ; LENGTH: 1209 base pairs
50 ; TYPE: nucleic acid
51 ; STRANDEDNESS: double
52 ; TOPOLOGY: circular
53 ; MOLECULE TYPE: DNA (genomic)
54 ; HYPOTHETICAL: NO
55 ; ANTI-SENSE: NO
56 ;
57 ; ORIGINAL SOURCE:
58 ; ORGANISM: Enterococcus faecium
59 ;
60 ; FEATURE:
61 ; NAME/KEY: misc feature
62 ; LOCATION: (8) LOCATION 1...1209
63 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1706:
64
65 US-09-107-532A-1706

```

[illegible]

Db 566 TCAAGATTACAGAGCGTTGGGAAAAACGAT 599

RESULT 10  
US-09-252-991A-12165  
Sequence 12165, Application US/09252991A

Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 12165  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12165

Query Match 0.9%; Score 39.6; DB 4; Length 1509;  
Best Local Similarity 45.3%; Pred. No. 0.23;  
Matches 144; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 3638 TTAGCGAAGAGTCGCGATCGTCCGACCTGATTCACAGCTCTGAAGACCGTATCGGCG 3697  
DB 482 TCACCCCGATGAAGTGCAGCCGAGGCGCTGCTGTGTGTCGCGGACATGCTCCGCGCG 541  
QY 3698 AAGTCGAGAAAGATATGCGAGTTGGTGTCAAAAACAAGAGATGAGACGACGACG 3757  
DB 542 ATGTGAGATTCCGTGGATGCGATGACATGACAGGTGCTATGAAAGCCACGCGCTGC 601  
QY 3758 CCTACCCCTGTCCGCGTCAACAGGTTTATGAAAAATGTTGCGCATCAAGCTGAGGCGA 3817  
DB 602 CCACCTACTTCTCTGCGACGTCGTGTCAGACACCTGATGAGCATCAACCACTCTGC 661  
QY 3818 TGACAAATCCGAGCAATATGATTTAAGTGATGATGAGTTGCTGAGCTGTCTTCC 3877  
DB 662 GCGCGAAGAGTGGTGGCGTCCGCGCCCAAGCTGATCAACTCAGAGATTTCCGCT 721  
QY 3878 TCGCGAGCCGTGAGATGATATCATGGGCTGCTGAGGCGTGAAGCGCTTCAAGCTGT 3937  
DB 722 GGGAGCAGCCGACGCTGTGTACATGCGCATGCGCAACCGGACAAAGAGAGCTGT 781  
QY 3938 ACTACCAAGAGCCCGCA 3955  
DB 782 CCAAGCCGCAAGATCCGA 799

RESULT 11  
US-09-252-991A-12425/c  
Sequence 12425, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 12425  
LENGTH: 1692

TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12425

Query Match 0.9%; Score 39.6; DB 4; Length 1692;  
Best Local Similarity 45.3%; Pred. No. 0.23;  
Matches 144; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 3638 TTAGCGAAGAGTCGCGATCGTCCGACCTGATTCACAGCTCTGAAGACCGTATCGGCG 3697  
DB 1124 TCACCCCGATGAAGTGCAGCCGAGGCGCTGCTGTGTGTCGCGGACATGCTCCGCGCG 1065  
QY 3698 AAGTCGAGAAAGATATGCGAGTTGGTGTCAAAAACAAGAGATGAGACGACGACG 3757  
DB 1064 ATGTGAGATTCCGTGGATGCGATGACATGACAGGTGCTGATGAAGCCGACGCGCTGC 1005  
QY 3758 CCTACCCCTGTCCGCGTCAACAGGTTTATGAAAAATGTTGCGCATCAAGCTGAGGCGA 3817  
DB 1004 CCACCTACTTCTCTGCGCAAGCTGTGTCAGACGACCTGATGGGATCAACCACTCTGC 945  
QY 3818 TGACAAATCCGAGCAATATGATTTCTAGGATCAGGTTGCTGAGCTGTCTTCC 3877  
DB 944 GCGCGAAGAGTGGTCCGCTCGCGCCCAAGCTGATCAAGCTTACAGATTTCCGCT 885  
QY 3878 TCGCGAGCCGTGAGATGATATCATGGGCTGCTGAGGCGTGAAGCGCTTCAAGCTGT 3937  
DB 884 GGGAGCAGCCGACGCTGTGTACATGCGCATGCTGTGCGCAACCGGACAAAGAGAGCTGT 825  
QY 3938 ACTACCAAGAGCCCGCA 3955  
DB 824 CCAAGCCGCAAGATCCGA 807

RESULT 12  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHREIFLINGER, F.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESS: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:



```

; LENGTH: 36093
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: m1ec_feature
; LOCATION: (1)..(36093)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14664

```

```

Query Match      0.9%; Score 38.6; DB 4; Length 36093;
Best Local Similarity 52.9%; Pred. No. 5.3;
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

```

QY      24  GAGGAATAGCCCGTCGAGAAATCATTAACCGGCTCAATAAGACTTACAACTGGGCT 83
Db      16851 GATGAGACAGCCTATCAAGGTGACTGAGTGGCAGACATACACTTACGACTCGGTAT 16910

QY      84  CCACTGTGTGCGAGACACAACTTCACCCCGCCGGAAGAGAGCTGGCCGAGAGTGA 143
Db      16911 CCACTCGGGGGCCAAACCTGCTGCTCCCTCCGTCAGCAGCAAGGCATCATGAGAGAGA 16970

QY      144  CGAGGATTTGGTGGCCATGCAAGATTTACAGAGCC 180
Db      16971 TGAGGCTGCGGGGCGCCAGTACAGCTCAAGAAAAC 17007

```

Search completed: August 10, 2005, 13:52:33  
 Job time : 629 secs

**This Page Blank (uspto)**



APPLICATION NUMBER: US 08/811,583  
 FILING DATE: 05-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
   NAME: Haley, James F.  
   REGISTRATION NUMBER: 27,794  
   REFERENCE/DOCKET NUMBER: MPG-1  
   TELECOMMUNICATION INFORMATION:  
     TELEPHONE: 212-596-9000  
     TELEFAX: 212-596-9090  
 INFORMATION FOR SEQ ID NO: 1:  
   SEQUENCE CHARACTERISTICS:  
     LENGTH: 3731 base pairs  
     TYPE: nucleic acid  
     STRANDEDNESS: single  
     TOPOLOGY: linear  
   MOLECULE TYPE: cDNA  
   ORIGINAL SOURCE:  
     ORGANISM: Tomato  
   FEATURE:  
     NAME/KEY: CDS  
     LOCATION: 194..3535  
   SEQUENCE DESCRIPTION: SEQ ID NO: 1:

	Query Match	1.2%	Score 49	DB 9	Length 3731	
	Best Local Similarity	51.6%	Pred. No. 0.00041			
	Matches 112	Conservative 0	Mismatches 105	Indels 0	Gaps 0	
QY	2850	TACACTCCTATCGGACTGTGATGTCCTCGTGGCGGATCCCGAGCCATTCCCTAGTGA	2909			
Db	2425	TAATTTCAATTCGGAAGGAAATGTGTGTGTGCAGAAAAATATCATCTTGACCTCGTGA	2484			
QY	2910	TATCAACAGGGTTGGAGCAGCTTCACGCCAGAGTCCACAGCTGCACAGATGTAAATCAT	2965			
Db	2485	TATTCGCTGTTTTAAAGCTGTAAATGTTTCGAGCGCTGCACACATGTAGATTTGTGTTG	2544			
QY	2970	CTTCTACTAAGAGAGATGTACCGCTTCTACAGAACTATCTGGTGAGACTACGACGG	3029			
Db	2545	ATTCCCTCAGAAAAGAAAAAGACCTCATCCGAAATGATGTTCTGGGAATGATTTGGATGG	2604			
QY	3030	CGATATGGCCCTGGGCTCTCGTGGGATCCGGAATCGTC	3066			
Db	2605	GGATATCTACTTTTGTCTGGATCGGAATCAAGACATGATC	2641			

```

# RESULT 2
# US-10-156-761-7161
# Sequence 7161, Application US/10156761
# Publication No. US20030119018A1
# GENERAL INFORMATION:
# APPLICANT: OMURA, SATOSHI
# APPLICANT: IKEDA, HARUO
# APPLICANT: ISHIKAWA, JUN
# APPLICANT: HORIKAWA, HIROSHI
# APPLICANT: SHIDA, TADAYOSHI
# APPLICANT: SAKAKI, YOSHIYUKI
# APPLICANT: HATTORI, MASAHIRA
# TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
# FILE REFERENCE: 249-262
# CURRENT APPLICATION NUMBER: US/10/156,761
# CURRENT FILING DATE: 2002-05-29
# PRIOR APPLICATION NUMBER: JP 2001-204089
# PRIOR FILING DATE: 2001-05-30
# PRIOR APPLICATION NUMBER: JP 2001-272697
# PRIOR FILING DATE: 2001-08-02
# NUMBER OF SEQ ID NOS: 15109
# SEQ ID NO 7161
# LENGTH: 1164
# TYPE: DNA
# ORGANISM: Streptomyces avermitilis
# FEATURE:
# NAME/KEY: CDS
# LOCATION: (1)..(1164)

```

Query Match	1.1%	Score 47	DB 15	Length 1164
Best Local Similarity	46.9%	Fred. No. 0.00096		
Match 146	Conservative	0	Mismatches 165	Indels 0
			Gaps	0
QY	AGCCCGCTGAGAGAAATCATATAACCGGCTCATATAAGCATCAACCTGGGCTCCAGTGT	90		
Db	16 ACCCCGAGAGACAGTTTACCTTGGGCTGTGGACCTGTGGCTGGCAGAGGAAGGACCCG	75		
QY	91 GTGCGAGACACAACTTTCACCCCCCAACCGCGGAAGAGCTGGCCGAGATGACGAGAT	150		
Db	76 TTGCGCGACGCGACCCCGCGGCGCTTGAGACCCGGTCGAGACGGTGCACGCGCTCGCGGG	135		
QY	151 TTGCGTGCATGACANATCTACAGAGCCCTGAACTTCTCTACTGCGGAAAGATAC	210		
Db	136 CTGGGTGCCACGAGATGATCTTCCACGACGACGACTGATTCCTTCGTGGTCTTCGGAC	195		
QY	211 TCCCTGAAACGAGGAGAGCCAACTTCTTCATCGAGGACCAAGCTGCAGCTCGAATCG	270		
Db	196 ACCGAGGCGAGTGCACATCAAGGCGTTTCGGGACGAGGCTTCGAAGCCACCGGATAGGCC	255		
QY	271 GTGCCCAAAGCCACAGCCGACCTTGACACGCTTCCGTGGTCCAAAGAACTTCCCGCGCC	330		
Db	256 GTGCGAATGGCGACACCAACCTCTTCACGACCCCGTCTTCAAGGAGCGGGGCGTTTACCC	315		
QY	331 GCTACTGCGCG	341		
Db	316 GCCAAGACCG	326		

```

RESULT 3
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMTURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
; US-10-156-761-1

Query Match      1.1%; Score 47; DB 15; Length 9025608;
Best Local Similarity 46.9%; Pred. No. 0.17;
Matches 146; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY      31      AGCCCCCTCGAGGAATCATTAACCGGCTCAATTAACAGCTATCAACCTGGGCTTCACAGT 90
            |||||
DB      8551009  ACCCCCGAGGACAGCTTCACTTCGGCTGTGAGACCTCGGCTGCGAGGAGGACCCG 8551068
            |||||

QY      91      GTGCGACAGCACTTCACACCCCCACCGCCGCGAGAGAGCTGGCCGAGAGTGAACAGAT 150
            |||||
DB      8551069  TTTCGGCAGCGCAACCCGCGCGGCTTGAGACCCGGTGCAGACGGTGCAGCCCTTCGCGCGG 8551128
            |||||

```



Qy	151	TTGGTGGCCATGACAAAGATCTACAGAGCCCTGAACTTTCCTACCTGCGGAAAGATGAC	210
Db	8551129	CTGGGTGCCACGAGAGTACCTTCCACGACAGACCTGATGCCCTTCGGGTCTCGGAC	8551188
Qy	211	TTCCCTGAAACACGAGCAGAAGCCAACTTTCCTACATCGAGGCCCAAGCTGCGAGCTCGAACTGG	270
Db	8551189	ACCGAGGCGGAGTGCACATCAAGCGCTTCCGGCAGAGCCCTTCGAGAGCACCAGGCAATGGCC	8551248
Qy	271	GTGCCCCAAAGTCCACGCGCGACCTTGACACGCTTCCGTGTCGAAGAACTTCCCGCGCC	330
Db	8551249	GTGCGGATGGCGACACCAACCTCTTCACGACCCCGTCTTCAAGAGAGGGGGCTTACCC	8551308
Qy	331	GCTACTGCCCCG	341
Db	8551309	GCCACAGCACCG	8551319

RESULT 4  
US-09-918-995-3407  
; Sequence 3407, Application US/09918995  
; Publication No. US20030073623A1

```

? TITLE OF INVENTION: FROM NUCLEIC ACID SEQUENCES OBTAINED
?
? TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
?
? FILE REFERENCE: 20411-756
?
? CURRENT APPLICATION NUMBER: US/09/918,995
?
? CURRENT FILING DATE: 2001-07-30
?
? PRIOR APPLICATION NUMBER: US/09/235,076
?
? PRIOR FILING DATE: 1999-01-20
?
? NUMBER OF SEQ ID NOS: 36054
?
? SOFTWARE: FastSeq for Windows Version 3.0
?
? SEQ ID NO 3407
?
? LENGTH: 495
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? FEATURE:
?
? NAME/KEY: misc_feature
?
? LOCATION: (1)..(495)
?
? OTHER INFORMATION: n = A,T,C or G
?
? US-09-918-995-3407

```

Query Match	1.1%	Score 45.6;	DB 10;	Length 495;
Best Local Similarity	51.5%	Pred. No. 0.0017;		
Matches 105; Conservative	0;	Mismatches 99;	Indels 0;	Gaps 0;

[illegible]

RESULT 5 155-158/C  
US-10-123-155-158/C  
; Sequence 358, Application US/10123155  
; Publication No. US20030068794N1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresani, Maureen  
; APPLICANT: DeFors, Laura  
; APPLICANT: Dennoyers, Luc

```

APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 358
LENGTH: 1049
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-155-358

```

Query Match	1.1%;	Score 44.8;	DB 14;	Length 1049;
Best Local Similarity	5.2%;	Pred. No. 0.0049;		
Matches	42;	Conservative 229;	Mismatches 531;	Indels 2;
				Gaps 1;

[illegible]

QY 871 AATCAGACACTACCCAGAGTAGTTTGGAGGCTCTCTCTTCACAGCCAGAGAGACGG 930  
D 348 S.H....S...D..MTHC..MT..N.S.H...NN.NM.WY..BD..SNSD.Y.BM..A..289  
QY 931 CCTGTGATGCCAGCGCTTTGGAGCTGGACACTTGATGCTCTCCTAGCAAGAGAGA 990  
D 288 TH..TNB.....SNMG..TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R...229  
QY 991 ACAACCAAGTCCCATAGATATAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

RESULT 6  
US-10-146-731-358/C

Sequence 358, Application US/10146731  
Publication No. US20030129692A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C123  
CURRENT APPLICATION NUMBER: US/10/146,731  
CURRENT FILING DATE: 2002-05-15  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-146-731-358

Query Match 1.1%; Score 44.8; DB 15; Length 1049;  
Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGGAACCGAGGAGCAACTTCTTCATCGAGCCAAAGTCGAGCTGCAACTGG 270  
D 1006 YMY....MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMWS...947  
QY 271 GTGCCCAAGGCCGAGCGGACCTGACAGCGCTTCGTGTCACAGAACTCCCGCGGCC 330  
D 946 S.S.N..B..C..MNY..G.N.DM.Y.H..MTB..T.BW.ABT.HNAHRE..THRGH..887  
QY 331 GCTACTGCCGCGCAACATGGGCACTGTGTTGCTCGAGGTGCTTAATAGGTTT 390  
D 886 S..Y..RCM.MTMG.DR.RMBHM.R.DSTAKKBK...SBS.S.S...NA.H..GAR.H.S 827  
QY 391 ATGCGACCTCCCATATACACACGAGTCGAACGTTTGGACAACTCTAAGGCGCCCAAGT 450  
D 826 .BS.CMPTC.CBGBHAT.R..AB.ADNBWB..WBTGAG..YND..KM.NN.BN...767  
QY 451 GGCCTGAGCGCGCCCACTTACCAACGCAAGAGTAGAGCGCGCCCAATGTCAC 510  
D 766 SAM..K..MNS.H.RY...TH...RMA.ST..NM...NM--..SYGNS.Y..BAA..N 709

QY 511 TTGCGATGCCGCCAAACGCTGTGACTCGCTTCCACAGCTCTCTATTCAGGC 570  
D 708 DS.H.A..NM.G..MMWS..SM.CNMT.S.NM.N..KCH.BCS...S.SNWS.H.....M...649  
QY 571 GCGCGATACCCCTAAAGTTCCCGATCCAGTGAATACCGGTTCCAAAGACCATCTCTC 630  
D 648 NM....RYNHC.YW.BH.DNCY...AY.S.S.KAYSASS.NHNHMK.M..BM.NMA.N..589  
QY 631 GAGATGAGATCTCAATCATGACACCAAGCGGCGCAAGCTGTCTGATTAATGTT 690  
D 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S..529  
QY 691 GCGGTCGCGCGCGCGCGCGCGCGCTTATGAGCGCTTGGACAGAGTACCGACTGA 750  
D 528 .NCS.N.GM..S.D..HSSWB...SNWS.H.A.CRMRS.NBSK.ST.NMN.YGSYTRMH 469  
QY 751 AGGCATGCCAATACGAGAGTCCCGGCGACAGGTCATGAGAGCGGACCGAGTGAT 810  
D 468 RY.RD...B...RS.BSAYTNSG..CB.SSHCS.S.MNBS.H.BM.YM..M.KS.MT.M..409  
QY 811 TCCTTGATATCATCTCAAGGACCTCTATGTTGAGTGTCTTCAGCGCTTCCGCTCAC 870  
D 408 .MAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYB...N..349  
QY 871 AATCAGACACTACCCAGAGTAGTTTGGAGCTCTCTCTTCACAGCCAGAGAGACGG 930  
D 348 S.H....S...D..MTHC..MT..N.S.H...NN.NM.WY..BD..SNSD.Y.BM..A..289  
QY 931 CCTGTGATGCCAGCGCTTTGGAGCTGGACACTTGATGATGATCTCCTAGCAAGAGAGA 990  
D 288 TH..TNB.....SNMG..TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R...229  
QY 991 ACAACCAAGTCCCATAGATATAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

RESULT 7  
US-10-140-472-358/C

Sequence 358, Application US/10140472  
Publication No. US2003013888A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C168  
CURRENT APPLICATION NUMBER: US/10/140,472  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-472-358

Query Match

1.1%; Score 44.8; DB 15; Length 1049;

	Beat	Local	Similarity	5.2%	Pred	No.	0.0049	
	Matches	42;	Conservative	229;	Mismatches	531;	Indels	2;
								Gaps
								1
Qy	211	TCCTGTAACCAAGGAGAAAGCCAACTTTCTTCATCGAGGCCAAAGCTGGCAGCTGCAACTGG	270					
Db	1006	YMY...MSM...M.....BHEM,HK.Y,DS.R,T,M,N,AMTRHAKV,BAMVS...	947					
Qy	271	GTGCCCAAGCCGACGCGCCAGCCCTGACAGCTTCCTGCTCCAAAGAACTCCCGCGCC	330					
Db	946	S.S.N'.B'.C'.WHY...G.N.DM.Y.H'.MTB..T.BW.ABT.HMAHRB..THRGH.	887					
Qy	331	GCTACTGCGGCCCAACATGAGCGCATTTGCGAGACTGTGTCTGCTGAGATGCTTAATAGATT	390					
Db	886	S..Y.RCM,MTMG.DR.RWBHW.R.DSTRKKK...SBS.S.S...NA.H.GAR.H.S	827					
Qy	391	ATGCCACCTCCCAATACACACACAGGAGCGAAGTGGAGTGGCAGAACTCTAAGCGGCCAAGT	450					
Db	826	.BS.CMDTC.CBGBAHAT.R..AB.ADBWMB.WBTHAG..YND..KM.NN.BN...	767					
Qy	451	GCGCTGAGCGCCGCAACCTCTTACCAACACCAAGGAGATGAGACCGGCAATGTACT	510					
Db	766	SAM..K..MNSS.H.RY...TH...RMA.SY..NNM...NM--..SYSGN.S.Y..BAA..N	709					
Qy	511	TTGCGTATCGGCCCAAGCGCTGTGTAAGTGGCTGGCCAGCGTCTCTATTCAGCGC	570					
Db	708	DS.H.A..NM.G..MMMS..SM.CNNT.S.NM.N..KCH.BCS..S.SNMS.H...M..	649					
Qy	571	GCGGCGATACCCCTTAAAGTTCCCGCAGTCCAGTAATCCGATTCCAAAGCAACTCTCTC	630					
Db	648	NM...RYNH.C.YW.BH.DNCY...AY.S.S.KAYSASS.HNHKK.M..BM.NMA..N..	589					
Qy	631	GAGAGTGGAATCTCAATCAGTGCACCAAGCGGCCAAAGGCGAACAAGCTGTGATATGTT	690					
Db	588	KDA.C.S..ROSNSS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S..	529					
Qy	691	GCGCGTGGCGCGCGCGCGCGCGCTGCTATTGCGAGCGCTTGGACAAGTACCGACTGCA	750					
Db	528	.NCS.N.GM..S.D..HSSMB...SNMS.H.A.CMKRS.NBSK..ST.MNM.YGSYTRMH	469					
Qy	751	AGGCATGCCAATACAGAGATCCACGCGCGACAGGTCAATAGACGCGACCGACAGGTGAT	810					
Db	468	RY.RD...B...RS.BSAVTMSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS..NT..M..	409					
Qy	811	TCCTTTGATACATCTCAAGGCACTTCTTATGTTGAGTGTTCAGAGCGCTTGGCGTAC	870					
Db	408	.NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S..NKSTYR...N..	349					
Qy	871	AATGAGACATACCCAGAGTAGTATTGAGGCTCTCTCTTACACACCCAGAGAGAAGGG	930					
Db	348	S.H...S...D..MTHC..MT..N.S.H...MM.NM.WY..BD..SNSD.Y..BM..A..	289					
Qy	931	CCTGTGATGCCACGCTTTTGAGGCTGGACACTGATGATGACTCTCTAGCAAGAAGA	990					
Db	288	TH.TVB...SNNG.TG...TNRGY.GNGS.H...N.NN.HH...MT.KXNR.R...	229					
Qy	991	ACAACCAAGTCCCATATGATTAAC	1014					
Db	228	A.AS..BA.BTABNNHM.S.BM.M	205					

RESULT 8  
US-10-141-761-358/c  
; Sequence 358, Application US/10141761  
; Publication No. US20030148432A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Wautreen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: prt
; ORGANISM: Homo Sapien
; US-10-141-761-358

```

Query Match	1.1*	Score 44.8	DB 15	Length 1049
Best Local Similarity	5.2*	Pred. No. 0.0049		
Matches	42	Conservative 229	Mismatches 531	Indels 2
				Gaps 1*
QY	211	TCGCTGAACCGAGGAAGCAACTCTTCTTCATGAGGCGCAAGAGCTGCGAGCTCGAATCG	270	
DB	1006	YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTBMAKB.BAMS...947		
QY	271	GTGCGCAAGCGCCAGCGGACCTTGACAGGCTTCGGTGTCCAGAAAGAACTCCCGCGCC	330	
DB	946	S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHRB..THRGH..887		
QY	331	GCTACTGCGCGCCACACATGAGGCAATGCGAGCTGTGTGCTCGAGCTGCTTAATAGTTT	390	
DB	886	S..Y.RCM.MTMG..DR.RWBHW.R.DSTAKBK...SBS.S.S...NA.H..GAR.H.S.827		
QY	391	ATGCGACCTCCCATTAACACACAGGTGAGCGTTTGGCAGAGCTTAAGCGGCCAAGT	450	
DB	826	BS.CMDTC.CBGBBHT.R..AB.ADNBWB.BETHAG..YVND..KM.NN..BN...767		
QY	451	GGCGTGAAGCCGCCCACTCTTACCAACCAACGCAAGATGAGCCGCCAATGCTACT	510	
DB	766	SAM..K..MNS.H.RV..TH...RMA.SV..NNM..NM--..SYSNS.Y..BAA..N.709		
QY	511	TTGCGTATCCGCCCAAGCGTCTGTTGACTGCTCTGCCACAGGCTCTCTTAATTCAGCG	570	
DB	708	DS.H.A..NM.G..MMS.SM.CMNT.S.NM.N..KCH.BCS...S.SNMS.H...M...649		
QY	571	GGCGGATACCCCTTAAGTTCCCGGATCCGATCCAGTAATCCGGTTCCAAGCAATCTCTC	630	
DB	648	NM...RNNH.YW.BH.DNCY...AY.S.S.KAYSASS.HNNHK.M..BM.NMA..N.589		
QY	631	GAGATGAGATTCATCAATCACTGACACCAAGCGGCGCAAGGCAAGCTGTCTGATTAATGT	690	
DB	588	KDA.C.S..RBSNS.H.B..MD...TASD..H.YNS.H.RV..T....SCN.A.S.529		
QY	691	CGCGTGCAGCGCGCGCGCGCTGATTTGCGAGCGCTTTGACAAAGTACGACTCGA	750	
DB	528	NCS.N.GM..S.D...HSSMB...SNMS.H.A..CRMGS..NBSK.ST.MNM.YGSYTRMH	469	
QY	751	AGGATGCGAATAGAGAGATCCACGCGGACAGAGTATATAGACAGCGGACAGAGTGAT	810	
DB	468	RY.RD...B...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H..BM.YW..M.KS..NT..M.409		
QY	811	TCTTTGATACATCTCAAGGCACTTCTATATGTTGAGTGCTTTCAGGCTGCGCTCAG	870	
DB	408	NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSYTB...N.349		
QY	871	AATGAGAGACTACCCAGAGTAGTTTGGAGGCTCTCTTCACAGCCGAGAGAGAGCGG	930	
DB	348	S.H...S...D..MTIC..MT..N.S.H...NM.NM.YV...BD..SNSD.Y..BM..A.289		

QY 931 CCTGTGATGCCAGGCTTTGAGCTGGACCTTGATGATCTCTTACCAAGAGA 990  
D 288 TH.TNB.....SNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R... 229  
QY 991 ACAACCAAGTCCCAATAGATTAAC 1014  
D 228 A.AS..BA.BTANNHM.S.BM.M 205

## RESULT 9

US-10-142-885-358/c

; Sequence 358, Application US/10142885  
; Publication No. US20030157604A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C248  
; CURRENT APPLICATION NUMBER: US/10/142,885  
; CURRENT FILING DATE: 2002-05-10  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-142-885-358

Query Match 1.1%, Score 44.8; DB 16; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGACCAAGCAAGCACTTTTCATCGAGCCAAAGCTGCGAGCTGCACTGG 270  
D 1006 YMY...MSM...M.....BHM.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMMS... 947  
QY 271 GTGGCCAAAGCCGCGGCGGACCGCTGACACGCTCCGCTGCAAGAACTCCCGGCGC 330  
D 946 S.S.N..B..C..WHY..G.N.DM.Y.H..MTB..T.BW.ABT.HVAHNB..THRGH. 887  
QY 331 GCTACTCGGCGCAACATGGGCACTGTGTGCTGCGAGTGTCTTAATAGGTTT 390  
D 886 S..Y.RCM.MTMG.DR.RWBHW.R.DSTAKBK...SBS.S.S...NA.H.GAR.H.S 827  
QY 391 ATGCCACCTCCCAATTAACACACGAGTGAAGCTTTGCGAAGCTTAAGCGGCCAAGT 450  
D 826 .BS.CMDTC.CVBABHAT.R..AB.ADNBWB.WBTGAG..YVDD...KM.NN.BN... 767  
QY 451 GGCGTAGCGCGCCCAACTCTACCAACAAAGCAAGATGAGCCCGCCCAATGTCAT 510  
D 766 SAM..K..MNS.H.RY...TH...RMA.ST..NM..NM--..SYGNS.Y..BAA..N 709  
QY 511 TTGCGTATCGCCCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570  
D 708 DS.H.A..NM.G..MMS.SM.CMNT.S.NM.N..KCH.BCS...S.SMS.H....M. 649

QY 571 GCGCGATACCCCTTAAGTTCCCGGATCCAGTGAATACCGGTTCCAAAGCATTCTTC 630  
D 648 NM...RYNHC.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHNK.M..BM.NNA.N. 589  
QY 631 GAGATGAGAAATCTCAATCAGTGCACCAAGCGGCCAAGGCGCAAGCTGTGTAATGTT 690  
D 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNS.H.RY..T....SCN.A.S. 529  
QY 691 GCGCGTCCGCGCGCGCGCGCGCGCGCTTAATGCGAGCGCTTTGCAAGATGACCTGCA 750  
D 528 .NCS.N.GM..S.D..HSMB...SNMS.H.A.CMRGS.NBSK.ST.NMM.YGYTRM 469  
QY 751 AGGATGCCAATACGAGATGCCAGCGGACAGCTCATAGACGACGCGAGTGGAT 810  
D 468 RY.RD...B...RS.BSAYTNSG.CB.SSHCS.S.MNS.H.BM.YM..M.KS..NT.M. 409  
QY 811 TCCTTGATCATCTCAAGCACTTCTTAATGTTGAGTGTCTTCAAGCGTCCGTCAC 870  
D 408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYB...N. 349  
QY 871 AATCAGACACTACCCAGAGTATTTTGAAGCTCTCTTCAAGCCGCAAGAAAGCG 930  
D 348 S.H...S..D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCAGGCTTTGAGCTGGACCTTGATGATGATCTCTTACCAAGAGA 990  
D 288 TH.TNB.....SNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R... 229  
QY 991 ACAACCAAGTCCCAATAGATTAAC 1014  
D 228 A.AS..BA.BTANNHM.S.BM.M 205

## RESULT 10

US-10-158-790-358/c

; Sequence 358, Application US/10158790  
; Publication No. US20030180879A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C448  
; CURRENT APPLICATION NUMBER: US/10/158,790  
; CURRENT FILING DATE: 2002-05-30  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-158-790-358

Query Match 1.1%, Score 44.8; DB 16; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGACCAAGCAAGCACTTTTCATCGAGCCAAAGCTGCGAGCTGCAACTGG 270

```

Db      1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMMS... 947
Qy      271 GTGCCCCAAAGCCGCGGAGCCCTGACAGCGCTCCGCTGTCACAGAACTCCCGCCGCC 330
Db      946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHKB..THRGH.. 887
Qy      331 GCTACTGCGGCGCAACATGGGCAATGGAGACTGTGTTGCTGAGAGTCTTAATAGTTT 390
Db      886 S..Y.RCM.MTMG.DR.RMBHW.R.DSTAKBK..SBS.S.S...NA.H.GAR.H.S 827
Qy      391 ATGCACCTCCCAATACACACACAGTCGACGTTTGACAGAACTTAAGCGGCCCAAGT 450
Db      826 .BS.CMDTC.CBAGBHAT.R..AB.ADNBWB.WBTHAG..YND..KM.NN.BN... 767
Qy      451 GGCCTGAGCCGCCCACTTACACACCAAGCCCAAGATGAGCCGCCCAATGTACT 510
Db      766 SAM..K..MNS.H.RY...TH...RMA.SY..NNM...NM--..SYGNS.Y..BAA..N 709
Qy      511 TTGCGTATCGCCCAAGCGCTGTGACTGCTGCTGTCACAGGCTCTTATTCAGGC 570
Db      708 DS.H.A..NM.G..MMS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H...M.. 649
Qy      571 GCGCGATACCCCTAAAGTTCCCGATCCAGTGAATACCGGTTCCAAAGCAATCTCTC 630
Db      648 NM...RYNHC.YW.BH.DNCY..AY.S.S.KAYSASS.HNNHK.M..BM.NMA.N.. 589
Qy      631 GAGAGTGAATCTCAATCACTGACGACCAAGCGGCCCAAGGCAAGCTGTGATPATTT 690
Db      588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S.. 529
Qy      691 GCCGTCGCGCGCGCGCGCGCTGCTATGCGAGCGTTTGACCAAGTACGACTGA 750
Db      528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRMRS.NBSK.ST.MNM.YGSTRMH 469
Qy      751 AGGCATGCCAATACAGAGATCCACGCGGACAGGTCATAGACGAGCAGCATGTGAT 810
Db      468 RY.RD...B...RS.BSAYTNSG.CB.SSHCS.S.NMBS.H.BM.YM..M.KS.NT.M.. 409
Qy      811 TCCTTTGATACATCTCAAGGCACTTCTTATGTTGAGTGTCTTCAAGCGCTTCCGCTAC 870
Db      408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM..SS.T.S.NKSTRB...N.. 349
Qy      871 AATCAGACACTACCCAGATAGTTTGAAGCTCCTCTTCAAGCCGACAGAGAAGCG 930
Db      348 S.H...S...D..MTHC..MT..N.S.H...NM.NM.WY..BD..SNSD.Y.BM..A.. 289
Qy      931 CCTGTGATGCGACGCTTTGAGGCTGACACTTGATGATCTCTTACAGCAAGAGA 990
Db      288 TH.TNB...SNMNG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNNR.R.. 229
Qy      991 ACAACCAAGTCCACATAGATAAC 1014
Db      228 A.AS..BA.BTABNNHM.S.BM.M 205

```

## RESULT 11

```

US-10-137-871-358/c
; Sequence 358, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria

```

```

; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C153
; CURRENT APPLICATION NUMBER: US/10/137, 871
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQUENCE ID NOS: 550
; NUMBER OF SEQUENCE ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-358

```

Query Match 1.1%, Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;

Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

```

Qy      211 TCCTGAACGAGGAGGAGGCACTTCTTATGAGGCGCAAGCTGCGAGCTGACTGG 270
Db      1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMMS... 947
Qy      271 GTGCCCCAAAGCCGCGGAGCCCTGACAGCGCTCCGCTGTCACAGAACTCCCGCCGCC 330
Db      946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHKB..THRGH.. 887
Qy      331 GCTACTGCGGCGCAACATGGGCAATGGAGACTGTGTTGCTGAGAGTCTTAATAGTTT 390
Db      886 S..Y.RCM.MTMG.DR.RMBHW.R.DSTAKBK..SBS.S.S...NA.H.GAR.H.S 827
Qy      391 ATGCACCTCCCAATACACACACAGTCGACGTTTGACAGAACTTAAGCGGCCCAAGT 450
Db      826 .BS.CMDTC.CBAGBHAT.R..AB.ADNBWB.WBTHAG..YND..KM.NN.BN... 767
Qy      451 GGCCTGAGCCGCCCACTTACACACCAAGCCCAAGATGAGCCGCCCAATGTACT 510
Db      766 SAM..K..MNS.H.RY...TH...RMA.SY..NNM...NM--..SYGNS.Y..BAA..N 709
Qy      511 TTGCGTATCGCCCAAGCGCTGTGACTGCTGCTGTCACAGGCTCTTATTCAGGC 570
Db      708 DS.H.A..NM.G..MMS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H...M.. 649
Qy      571 GCGCGATACCCCTAAAGTTCCCGATCCAGTGAATACCGGTTCCAAAGCAATCTCTC 630
Db      648 NM...RYNHC.YW.BH.DNCY..AY.S.S.KAYSASS.HNNHK.M..BM.NMA.N.. 589
Qy      631 GAGAGTGAATCTCAATCACTGACGACCAAGCGGCCCAAGGCAAGCTGTGATPATTT 690
Db      588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S.. 529
Qy      691 GCCGTCGCGCGCGCGCGCGCTGCTATGCGAGCGTTTGACCAAGTACGACTGA 750
Db      528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRMRS.NBSK.ST.MNM.YGSTRMH 469
Qy      751 AGGCATGCCAATACAGAGATCCACGCGGACAGGTCATAGACGAGCAGCATGTGAT 810
Db      468 RY.RD...B...RS.BSAYTNSG.CB.SSHCS.S.NMBS.H.BM.YM..M.KS.NT.M.. 409
Qy      811 TCCTTTGATACATCTCAAGGCACTTCTTATGTTGAGTGTCTTCAAGCGCTTCCGCTAC 870
Db      408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM..SS.T.S.NKSTRB...N.. 349
Qy      871 AATCAGACACTACCCAGATAGTTTGAAGCTCCTCTTCAAGCCGACAGAGAAGCG 930
Db      348 S.H...S...D..MTHC..MT..N.S.H...NM.NM.WY..BD..SNSD.Y.BM..A.. 289
Qy      931 CCTGTGATGCGACGCTTTGAGGCTGACACTTGATGATCTCTTACAGCAAGAGA 990
Db      288 TH.TNB...SNMNG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNNR.R.. 229

```

QY 991 ACAACCAAGTCCCATAGATAAC 1014  
Db 228 A.AS..BA.BTABNNHM.S.BM.M 205

## RESULT 12

US-10-140-923-358/c  
; Sequence 358, Application US/10140923  
; Publication No. US20030207355A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C18  
; CURRENT APPLICATION NUMBER: US/10/140, 923  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-923-358

Query Match 1.1%, Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCCTGAACGAGCAAGCACTTCTTCATCGAGCCAAAGCTGCGAGCTGCACTGG 270  
Db 1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTMHAAB.BAMMS... 947  
QY 271 GTGCCCAAGCCGACGCGGACCTGACAGCGCTTCGTTGTCGAAGAACCTCCCGCGCC 330  
Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHNB..THRGSH. 887  
QY 331 GCTACTCGCGCCCAACATGAGGCTTCAGACGTGTCTTCGCGAGGCTTAAAGTTT 390  
Db 886 S..Y.RCM.MTMG..DR.RMBHM.R.DSTAKBK...SBS.S.S...NA.H..GAR.H.S 827  
QY 391 ATGCCACCTCCCAATAACACACAGCTGCAAGCTTTGSCAGAACTCTTAAGGCGCCAAGT 450  
Db 826 .BS.CMPTC.CVBGBHAT.R..AB.ADNBWB.BMTGAG..YVDD...KM.NN.BN... 767  
QY 451 GGCCTGAGCGCCCAACCTCTTACACACAAAGCAAGTAGAGCCCGCAATGTCACT 510  
Db 766 SAM..K..MNS.S.H.RV...TH...RMA.SY...NNM...NM--..SYSGNS.Y..BAA..N 709  
QY 511 TTGCTGATCGCCCAACAGCTGCTGACTGCTCTGCCACAGGTCTCTTATTCAGCGC 570  
Db 708 DS.H.A..NM.G..MMS.SM.CMNT.S.NM.N..KCH.BCS...S.SNMS.H...M.. 649  
QY 571 GCGGCAATACCCGTAAGTTCCCGATCGAGTGAATGAGGTTCCAAAGCACTCTCTC 630  
Db 648 NM...RYNHC.YW.BH.DNCY...AY.S.S.KAYSASSS.HNNHKK.M..BM.NMA.N. 589

QY 631 GAGAGTGAATCTCATCATGACGACCAAGCGGCGCAAGGCAAGCTGTCTGATAATGTT 690  
Db 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T.....SCN.A.S. 529  
QY 691 GCGGCTGCG 750  
Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRMRS.NBSK..ST.MNM.YGSYTRM 469  
QY 751 AGCATGCGCAATAGAGATGCCAGGACAGTATGATAGAGCGGACGAGTGTGAT 810  
Db 468 RY.RD...B..RS.BSATYNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M. 409  
QY 811 TCTTGTAATCATCTCAAGCACTTCTTATGTTGCAAGTCTTCAAGCGTTCGCTCAC 870  
Db 408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB...N. 349  
QY 871 AATCAGAGCACTACCGAGTACTTTTGAGGCTCTTCAAGCCCGAGAGAGCGG 930  
Db 348 S.H...S...D..MTHC.MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCACGCTTTTGAGGCTGACACTGATGAGTCTCTCAAGCAAGAGA 990  
Db 288 TH.TNB...SNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R.. 229  
QY 991 ACAACCAAGTCCCATAGATAAC 1014  
Db 228 A.AS..BA.BTABNNHM.S.BM.M 205

## RESULT 13

US-10-141-756-358/c  
; Sequence 358, Application US/10141756  
; Publication No. US20030207355A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C200  
; CURRENT APPLICATION NUMBER: US/10/141, 756  
; CURRENT FILING DATE: 2002-05-08  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-141-756-358

Query Match 1.1%, Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCCTGAACGAGCAAGCACTTCTTCATCGAGCCAAAGCTGCGAGCTGCACTGG 270  
Db 1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTMHAAB.BAMMS... 947  
QY 271 GTGCCCAAGCCGACGCGGACCTGACAGCGTTCCGTTGTCGAAGAACCTCCCGCGCC 330

Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHNB..THRGH.. 887  
 Qy 331 GCTACTGCGGCGCAACATGGGCGATTCGAGACTGTGTTGCTCGAGTCTTAATAGTTT 390  
 Db 886 S..Y..RCM..MTMG..DR..RWBHM..R..DSTAKKBK...SBS..S.S...NA..H..GAR..H.S 827  
 Qy 391 ATGCACACTCCCAATTAACACACAGGATGTTGGACAGAACTTAAGGGGCGCAAGT 450  
 Db 826 ..BS..CMOTC..CBGBBHT..R..AB..ADNBWB..WBTHAG...YND..KM..NN..BN... 767  
 Qy 451 GGCCTGAGCCGCCCACTTACCAACCAACGCAAGATGAGCCGCCCAATGTCACT 510  
 Db 766 SAM..K..MNS..H..RY...TH...RMA..SY..NNM...NM--..SYSNS..Y..BAA..N 709  
 Qy 511 TTGCGTATCGGCGCAAGCGTCTGTAAGTGTGCTGCGACAGTCTCTTAATTCAGCGC 570  
 Db 708 DS..H..A..NM..G...MMMS..SM..CNMT..S..NM..N..KCH..BCS...S..SNMS..H...M... 649  
 Qy 571 GCGGCGATACCCCTAAAGTTCCCGGATCCAGATACCGGTTCCAAACGCAATCTCTC 630  
 Db 648 NM...RYNHC..Y..BH..DNCT...AY..S..S..KAYSASS..HNHNK..M..BM..NMA..N. 589  
 Qy 631 GAGAGTGAATCTCAATCACTGACGACCAAGCGGCGCAAGCGTCTGTATATGTT 690  
 Db 588 KDA..C..S...RDSNS..H..B..MD...TASD..H..YNS..H..RY..T...SCN..A..S. 529  
 Qy 691 GCGCGTCCCGCGCGCGCGCGCGTCTGATGAGCGCTTGGACAGATCCGACTGA 750  
 Db 528 ..NCS..N..GM..S..D...HSSMB...SNMS..H..A..CMRGS..NBSK..ST..MMN..YGSYTRMH 469  
 Qy 751 AGGATGCAATAGAGAGATCCCGCGCGGACAGTCAATAGACGCGGACCGAGTGTAT 810  
 Db 468 RY..RD...B...RS..BSAYTNSG..CB..SSHCS..S..MMBS..H..BM..YM..M..KS..NT..M. 409  
 Qy 811 TCCTTGATATCATCTCAAGGACCTTCTATGTTGATGTTGATGAGCGCTTGCCTGAC 870  
 Db 408 ..NAC..H..B..N..ND..S..N..SM..M..RCY..Y..M..SM..SS..T..S..NKSTYRB...N. 349  
 Qy 871 AATCAGAGACTACCCAGAGTAGTTTGAAGCTCTCTTCAACAGCCCAAGAGAGCGG 930  
 Db 348 S..H...S...D...MTHC...MT..N..S..H...NM..NM..WY..BD...SNSD..Y..BM..A. 289  
 Qy 931 CCTGTGATGCGACGCTTTTGAAGCTGACACTTGATGAGTCTCTTAAGCAAGGAAGA 990  
 Db 288 TH..TNB...SNNMG..TG...TNKGY..GNCS..H...N..NN..HH...MT..KNNR..R... 229  
 Qy 991 ACAACCAAGTCCACATAGATPAC 1014  
 Db 228 A..AS...BA..BTABNNHM..S..BM..M 205

RESULT 14  
 US-10-141-759-358/c  
 Sequence 358, Application US/10141759  
 Publication No. US20030207361A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: DeNovo, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gettleman, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William

APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE OF INVENTION: ACTS ENCODING THE SAME  
 FILE REFERENCE: P33081C197  
 CURRENT APPLICATION NUMBER: US/10/141.759  
 PRIOR FILING DATE: 2002-05-08  
 PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 358  
 LENGTH: 1049  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-10-141-759-358

Query Match 1.1%, Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred.No.0.0049; Mismatches 531; Indels 2; Gaps 1;

Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;  
 Qy 211 TCCCTGAACCGAGAGAGCAACTTCTCATCGAGCGCAAGAGCTGCGAGCTGCAACTGG 270  
 Db 1006 YMY...MSM...M.....BHM..HK..Y..DS..R..T..M..N..ATRNHAKB..BAMMS... 947  
 Qy 271 GTGCCAAAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330  
 Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHNB..THRGH.. 887  
 Qy 331 GCTACTGCGGCGCAACATGGGCGATTCGAGACTGTGTTGCTCGAGTCTTAATAGTTT 390  
 Db 886 S..Y..RCM..MTMG..DR..RWBHM..R..DSTAKKBK...SBS..S.S...NA..H..GAR..H.S 827  
 Qy 391 ATGCACACTCCCAATTAACACACAGGATGTTGGACAGAACTTAAGGGGCGCAAGT 450  
 Db 826 ..BS..CMOTC..CBGBBHT..R..AB..ADNBWB..WBTHAG...YND..KM..NN..BN... 767  
 Qy 451 GGCCTGAGCCGCCCACTTACCAACCAACGCAAGATGAGCCGCCCAATGTCACT 510  
 Db 766 SAM..K..MNS..H..RY...TH...RMA..SY..NNM...NM--..SYSNS..Y..BAA..N 709  
 Qy 511 TTGCGTATCGGCGCAAGCGTCTGTAAGTGTGCTGCGACAGTCTCTTAATTCAGCGC 570  
 Db 528 ..NCS..N..GM..S..D...HSSMB...SNMS..H..A..CMRGS..NBSK..ST..MMN..YGSYTRMH 469  
 Qy 751 AGGATGCAATAGAGAGATCCCGCGCGGACAGTCAATAGACGCGGACCGAGTGTAT 810  
 Db 468 RY..RD...B...RS..BSAYTNSG..CB..SSHCS..S..MMBS..H..BM..YM..M..KS..NT..M. 409  
 Qy 811 TCCTTGATATCATCTCAAGGACCTTCTATGTTGATGTTGATGAGCGCTTGCCTGAC 870  
 Db 408 ..NAC..H..B..N..ND..S..N..SM..M..RCY..Y..M..SM..SS..T..S..NKSTYRB...N. 349  
 Qy 871 AATCAGAGACTACCCAGAGTAGTTTGAAGCTCTCTTCAACAGCCCAAGAGAGCGG 930  
 Db 348 S..H...S...D...MTHC...MT..N..S..H...NM..NM..WY..BD...SNSD..Y..BM..A. 289  
 Qy 931 CCTGTGATGCGACGCTTTTGAAGCTGACACTTGATGAGTCTCTTAAGCAAGGAAGA 990  
 Db 288 TH..TNB...SNNMG..TG...TNKGY..GNCS..H...N..NN..HH...MT..KNNR..R... 229  
 Qy 991 ACAACCAAGTCCACATAGATPAC 1014  
 Db 228 A..AS...BA..BTABNNHM..S..BM..M 205





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 17:59:01 ; Search time 12576 Seconds

(without alignments)  
12730.464 Million cell updates/sec

Title: US-09-913-878a-1\_COPY\_2447\_6652

Perfect score: 4206  
Sequence: 1 atgacaccctactactcctag.....tcacagggaatgcgattat 4206Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hnc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsal:\*  
9: gb\_gsal2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	1.9	737	7	CF883391 trlC088xh
2	78	1.9	776	6	CB809520 trlC088xh
3	78	1.9	787	6	CB807399 trlC088xh
4	76.4	1.8	972	7	CF885460 trlC088xh
5	51.8	1.2	786	9	CG809201 FSAAE20TF
6	49	1.2	570	2	BE433646 EST400175
7	49	1.2	684	7	CR651141 11217 Stc
8	49	1.2	737	7	CR854274 15379 Stc
9	47	1.1	603	7	CV190148 vco01-6ms
10	47	1.1	925	9	CNS0091P
11	45.6	1.1	597	2	AM649811 EST328265
12	45	1.1	367	2	BF112386 EST439976
13	45	1.1	421	1	AI896465 EST65896
14	45	1.1	525	2	AM933692 EST59535
15	44.8	1.1	742	7	CF820058 EST679440
16	44.6	1.1	681	8	BH045290 RPCI-24-2
17	44	1.0	310	1	AA501295 vhe2d09.r
18	44	1.0	454	1	AA259012 z635G01.s
19	44	1.0	529	2	AW971719 EST83808
20	44	1.0	584	5	BP380240 BP380240
21	44	1.0	592	2	AM872991 hq2ia01.x
22	44	1.0	708	6	CA180649 SCACST315
23	44	1.0	720	9	AY409056 Homo sapi
24	44	1.0	720	9	AY409057 Pan trogl

25	44	1.0	958	1	AL559045
26	44	1.0	1844	3	CR609043 full-length
27	44	1.0	3022	3	AL833977 Homo sapi
28	43.6	1.0	1011	1	AL559276
29	43.4	1.0	521	5	BQ111608 EST597184
30	43.4	1.0	610	2	BB920043 EST423812
31	42.6	1.0	787	9	CNS01087 Drosophila
32	42.6	1.0	1206	9	CL979698 OeIFCC033
33	42.4	1.0	582	4	BI246074 IPL 68 D1
34	42.4	1.0	584	4	BP379014 BP379014
35	42.4	1.0	622	5	BU965311 sac08907.
36	42.4	1.0	641	6	CA257764 SCCCF110
37	42.4	1.0	681	1	A1125852 qe01f06.x
38	42.2	1.0	405	7	CN253957 BIP0019 B
39	41.8	1.0	966	9	CG006611 ZUAH588TV
40	41.2	1.0	555	6	CA866785 i472e02.x
41	41.2	1.0	563	5	BQ271751 i472h09.x
42	41.2	1.0	595	1	AI764981 wh57c04.x
43	41.2	1.0	651	5	BQ014248 UI-H-BD1-
44	41.2	1.0	658	8	B2336439 h234c12.b
45	41.2	1.0	715	7	CP865881 trlC003xm

## ALIGNMENTS

RESULT 1  
CF883391  
LOCUS CF883391 737 bp mRNA linear EST 31-OCT-2003DEFINITION trlC088xh09.b1 T.reesei mycelial culture, Version 6 October 2003  
ACCESSION CF883391  
VERSION CF883391  
KEYWORDS EST.ORGANISM Hypocrea jecorina (anamorph: Trichoderma reesei)  
Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

REFERENCE 1 (bases 1 to 737)

AUTHORS Diener, S.B., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,  
Teunissen, P.J.M., van Solingen, P., Dankmeyer, U., Mitchell, T.K.,  
Ward, M. and Dean, R.A.

QY 3889 GAGATGAATACATGCGATTGCTGAGGCGTACGACGCGCTTCAAGCTGTACTACACACAG 3948  
 DB 221 GCGTTGAGTACTGTTTTTGTCTGAAACATGCTTCACCACTTTCAAGAGTATCACACAGC 280  
 QY 3949 AGCCCCAAGTTGCTGTGCGATGCGGGGACAGCTCGCGATCAATTAAGGCGCAGATG 4008  
 DB 281 GCGCCCAAGTTTGTCTGAAACATGCGGAAAGCAGCTTGTCTACATCAAGGCGATGCGC 340  
 QY 4009 ACAGACGAGCCCGGTGAAGCGCGCCCGCTTGATGACCGCTTCATGTATGCGGCTTG 4068  
 DB 341 CGAAACGACGCTGGCGAGACGTCGAAAGTGTGTGATCCCGAGATGTGGGGGCTCTG 400  
 QY 4069 ATGCGCGATTAAGAGTTTACGAAAGCATATGTGCGCCAGCTCGAGGCGCATGTGATCGGAG 4128  
 DB 401 CGGCGCGACAAAGAGCTCATCCTCGCTGCGCATACAGAGAGGCTGCGCGGATTCG 460  
 QY 4129 TACCCTGATCCGAGGCTTATGAAGTCTGGGCGATGATGATTTTG 4174  
 DB 461 GAGAGTCACTGCTTGTGAGAGAGGTGTTGAGTTTGAAGTATGATG 506

RESULT 2  
 CB909520 776 bp mRNA linear EST 02-JUL-2003  
 LOCUS trico88xh09 T. reesei mycelial culture, Version 3 april Hypocrea  
 DEFINITION jecorina cDNA clone trico88xh09, mRNA sequence.  
 ACCESSION CB909520  
 VERSION CB909520.1 GI:30124181  
 KEYWORDS EST.  
 SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
 ORGANISM Hypocrea jecorina

REFERENCE  
 AUTHORS Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,  
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,  
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,  
 Olivares, H.A., Teunissen, P.J., Yao, J., and Ward, M.  
 Transcriptional regulation of biomass-degrading enzymes in the  
 filamentous fungus Trichoderma reesei  
 J. Biol. Chem. 278 (34), 31988-31997 (2003)

TITLE JOURNAL  
 MEDLINE 22803314  
 PUBMED 12788920  
 COMMENT Contact: Pamela K. Foreman  
 Genencor Intl.  
 925 Page Mill Road, Palo Alto, CA 94304, USA  
 Tel: (650) 846-7635  
 Fax: (650) 621-7817  
 Email: PForeman@genencor.com  
 Seq primer: LT-F1 primer.  
 Location/Qualifiers

FEATURES  
 source  
 1. 776  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="trico88xh09"  
 /dev\_stage="mycelia"  
 /note="Vector: PREP3; Site 1: Not I/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."

ORIGIN  
 Query Match 1.9%; Score 78; DB 6; Length 776;  
 Best Local Similarity 54.5%; Pred. No. 1.2e-10;  
 Matches 156; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 3889 GAGATGAATACATGCGATTGCTGAGGCGTACGACGCGCTTCAAGCTGTACTACACACAG 3948  
 DB 221 GCGTTGAGTACTGTTTTTGTCTGAAACATGCTTCACCACTTTCAAGAGTATCACACAGC 280  
 QY 3949 AGCCCCAAGTTGCTGTGCGATGCGGGGACAGCTCGCGATCAATTAAGGCGCAGATG 4008  
 DB 281 GCGCCCAAGTTTGTCTGAAACATGCGGAAAGCAGCTTGTCTACATCAAGGCGATGCGC 340  
 QY 4009 ATGCGCGATTAAGAGTTTACGAAAGCATATGTGCGCCAGCTCGAGGCGCATGTGATCGGAG 4128  
 DB 401 CGGCGCGACAAAGAGCTCATCCTCGCTGCGCATACAGAGAGGCTGCGCGGATTCG 460  
 QY 4129 TACCCTGATCCGAGGCTTATGAAGTCTGGGCGATGATGATTTTG 4174  
 DB 461 GAGAGTCACTGCTTGTGAGAGAGGTGTTGAGTTTGAAGTATGATG 506

DB 281 GCGCCCAAGTTTGTCTGAAACATGCGCGAAACACCTTGCCTACATCAAGGCGATGCGC 340  
 QY 4009 ACAGACGAGACCGCGTGAAGCGCGCCCGGCTTATGACCGCTTCAATGATGCGGCTTG 4068  
 DB 341 CGAAAGCAGCTGCGAGACGTCGAAAGTGTGTATTTCCGAGATGTGGGGGCTCTG 400  
 QY 4069 ATGCGCGATTAAGAGTTTACGAAAGCATATGTGCGCCAGCTCGAGGCGCATGTGATCGGAG 4128  
 DB 401 CGGCGCGACAAAGAGCTCATCCTCGCTGCGCATACAGAGAGGCTGCGCGGATTCG 460  
 QY 4129 TACCCTGATCCGAGGCTTATGAAGTCTGGGCGATGATGATTTTG 4174  
 DB 461 GAGAGTCACTGCTTGTGAGAGAGGTGTTGAGTTTGAAGTATGATG 506

RESULT 3  
 CB907399 787 bp mRNA linear EST 02-JUL-2003  
 LOCUS trico82xd19 T. reesei mycelial culture, Version 3 april Hypocrea  
 DEFINITION jecorina cDNA clone trico82xd19, mRNA sequence.  
 ACCESSION CB907399  
 VERSION CB907399.1 GI:30122057  
 KEYWORDS EST.  
 SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
 ORGANISM Hypocrea jecorina

REFERENCE  
 AUTHORS Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,  
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,  
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,  
 Olivares, H.A., Teunissen, P.J., Yao, J., and Ward, M.  
 Transcriptional regulation of biomass-degrading enzymes in the  
 filamentous fungus Trichoderma reesei  
 J. Biol. Chem. 278 (34), 31988-31997 (2003)

TITLE JOURNAL  
 MEDLINE 22803314  
 PUBMED 12788920  
 COMMENT Contact: Pamela K. Foreman  
 Genencor Intl.  
 925 Page Mill Road, Palo Alto, CA 94304, USA  
 Tel: (650) 846-7635  
 Fax: (650) 621-7817  
 Email: PForeman@genencor.com  
 Seq primer: LT-F1 primer.  
 Location/Qualifiers

FEATURES  
 source  
 1. 787  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="trico82xd19"  
 /dev\_stage="mycelia"  
 /note="Vector: PREP3; Site 1: Not I/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."

ORIGIN  
 Query Match 1.9%; Score 78; DB 6; Length 787;  
 Best Local Similarity 54.5%; Pred. No. 1.2e-10;  
 Matches 156; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 3889 GAGATGAATACATGCGATTGCTGAGGCGTACGACGCGCTTCAAGCTGTACTACACACAG 3948  
 DB 215 GCGTTGAGTACTGCGAGCTGCTGAAAGCTTGCAGACCTTCAAGAGTATCACACACAGC 274  
 QY 3949 AGCCCCAAGTTGCTGTGCGATGCGGGGACAGCTCGCGATCAATTAAGGCGCAGATG 4008  
 DB 275 GCGCCCAAGTTTGTCTGAAACATGCGCGAAAGCAGCTTGTCTACATCAAGGCGATGCGC 334  
 QY 4009 ACAGACGAGACCGGCTGAAGCGCGCCCGGCTTATGACCGCTTCAATGATGCGGCTTG 4068

Db 335 CGAAGCAGCGCTGGCGAGACGTCAAAGTCTTGATGCCGAGATGTGGGGCTCTG 394

Qy 4069 ATGCCGATAGAGTTTACAGACAGATGTGGCCAGCGCTGAGGGCGATGATCGGAG 4128

Db 395 CGGCGGAGCAAGAGCTATCACTCGCTGGCGATACAGAGGAGGCTGCCCGGAGATTG 454

Qy 4129 TACCCGTATCCGAGGTCTATGAAAGTCTGGCGCATGATTTTG 4174

Db 455 GAGAGTGCACCTGCTTGGAGAGGTGTGATGATTGACGATATG 500

RESULT 4  
CF885460 972 bp mRNA linear EST 31-OCT-2003  
LOCUS trico82xd19.b1 T.reesei mycelial culture, Version 6 October 2003  
DEFINITION Hypocrea jecorina cDNA clone trico82xd19, mRNA sequence.  
ACCESSION CF885460  
VERSION CF885460.1 GI:38140142  
KEYWORDS EST.  
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
ORGANISM Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.  
REFERENCE 1 (bases 1 to 972)  
Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D., Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K., Ward, M. and Dean, R.A.  
Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei  
FEBS Microbiol. Lett. 230 (2), 275-282 (2004)  
JOURNAL COMMENT Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA  
Tel: 919-513-0020  
Fax: 919-513-0024  
Email: ralph.dean@ncsu.edu  
Seq primer: LT-F1 primer  
FEATURES  
Location/Qualifiers  
source 1..972  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="OM6a"  
/db\_xref="taxon:51453"  
/clone="trico82xd19"  
/dev\_stage="mycelia"  
/clone\_1lb="T.reesei mycelial culture, Version 6 October 2003"  
/note="Vector: pREP3; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying carbon and nitrogen sources and concentrations."

ORIGIN  
Query Match 1.8%; Score 76.4; DB 7; Length 972;  
Best Local Similarity 54.2%; Pred. No.3.9e-10;  
Matches 155; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 3889 GAGATGAATACATGGGCTGCTGAGGCTGACACGCGCTTCAAGCTGTACTACCAAG 3948

Db 138 GCGTTGAGTAATCGGCACTGCTGAAAGCTTGCACATTTCAAGAAATATACACAGC 197

Qy 3949 AGCCCCAAGTTCGTGTGCAGATGCGGGGAGACAGCTCGGTCATTAAGCGCAGATG 4008

Db 198 GCGCCCAAGTTGTGTGAACATGCGCGAAGACGCTTGCTTACATAAAGCCAGTGGCG 257

Qy 4009 ACAGACAGACCCGCGTGAAGCGCCCGCGCTTGATGACCGGCTTCATGTATGCGGCTTG 4068

Db 258 CGAAACGACGCTGGCGAGACGTCGAAGGTGCTTGATTCGCGAGATGTGGGGCTCTG 317

Qy 4069 ATGCCGATTAAGATTATGAGAGCATATGTGGCGAGCTGAGAGGCGGATGATCGGAG 4128

Db 318 CGGCGGAGCAAGAGCTCATCACTCGCTGGCGATACAGAGGAGGCTGCCCGGAGATTG 377

Qy 4129 TACCCGTATCCGAGGTCTATGAAAGTCTGGCGCATGATTTTG 4174

Db 378 GAGAGTGCACCTGCTTGGAGAGGTGTGATGATTGACGATATG 423

RESULT 5  
CG809201/c 786 bp DNA linear GSS 13-NOV-2003  
LOCUS FSAAE20TR largeinsertGenomicLibrary Fusarium virguliforme genomic clone KMfV2C15, genomic survey sequence.  
DEFINITION CG809201  
ACCESSION CG809201  
VERSION CG809201.1 GI:38262675  
KEYWORDS GSS.  
SOURCE Fusarium virguliforme  
ORGANISM Fusarium virguliforme  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.  
REFERENCE 1 (bases 1 to 786)  
Meksem, K., Ishihara, H., Koo, H., Shultz, J., Ali, S., Iqbal, J., Lightfoot, D.A. and Town, C.D.  
End sequencing of BACs from a fingerprint physical map of the causative agent of soybean sudden death syndrome, Fusarium verticilliiforme  
Unpublished (2003)  
JOURNAL COMMENT Other GSSs: FSAAE20TR  
Contact: Chris Town and K. Meksem  
The Center of Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale and Plant Genomics, The Institute for Genomic Research  
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415, USA and 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 618 453 3103 and 301-838-3523  
Fax: 618 453-7457 and 301-838-0208  
Email: meksemk@uiu.edu; cdtown@igrr.org (URL: http://Fusariumvirguliforme.uiu.edu)  
Seq primer: CAGGAAACAGCTATGACC  
Class: BAC ends.  
FEATURES  
Location/Qualifiers  
source 1..786  
/organism="Fusarium virguliforme"  
/mol\_type="genomic DNA"  
/cultivar="Monticello"  
/db\_xref="taxon:232082"  
/clone="KMfV2C15"  
/clone\_1lb="largeinsertGenomicLibrary"  
/note="Organ: Hypophae; Vector: pINDIOBAC5; A single spore derived culture was used. Hypphae were grown in an incubator for four days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III. Large size DNA fragments were ligated in vector pINDIOBAC5 and electro-transformed into DH10B cells."

ORIGIN  
Query Match 1.2%; Score 51.8; DB 9; Length 786;  
Best Local Similarity 48.8%; Pred. No.0.0058;  
Matches 140; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 3883 GACCGTAGATGAATACATGGGCTGCTGAGGCTGACACGCGCTTCAAGCTGTACTAC 3942

Db 592 GACTCTCCCTCGAGCAATGGGAGCTCTCAAGCGCTCGACATGTTCAGAGCTCGGCTAC 533

Qy 3943 CACAAGAGCCCAAGTTCGTGTGCAGATGCGGGGAGACAGAGCTCGGTCATTAAGGCG 4002

Db 532 AAGTTGCGTACAGCATGTGTGCGGCTCTCGGCGACAGCATTTGCTGATTAAGGCC 473

Qy 4003 CAGATGACGAGACAGACCCGCGTGAAGCGCCCGCGCTTGATGACCGGCTTCATGTATGCG 4062

Db 472 ATGATGTCCCGCGGGGTGTGTGATGTGAGCGCGTGCCTGACAGCGCGAGATGTGAGC 413

Qy 4063 GCGTTAGTCCCGATTAAGATTATGAGAGCATATGTGGCGAGCTGAGAGGCGGATGGA 4122

Db 412 ATCTCCGCGCGACAAACAGCGATGCGGCGCTGACTGCTCCCTGCGCCAGGCTGACAT 353

QY 4123 TCGAGTACCTGATCCGAGGTCTATGAAGTCTGGCCGATGTA 4169  
DB 352 GACATGAGAGTGTGACGCGCTCGAGAGGTGATGATGATGTA 306

RESULT 6  
BE433646  
LOCUS BE433646 570 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST400175 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
ACCESSION BE433646  
VERSION BE433646  
KEYWORDS BE433646.1 GI:9431489  
SOURCE EST.  
ORGANISM Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE 1 (bases 1 to 570)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Roming,C.M.,  
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tanksey,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
COMMENT Unpublished (2000)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1..570  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA96"  
/db\_xref="taxon:4081"  
/clone="CLEG16J9"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: BluescriptSKmudapt; Site\_1: EcoRI;  
Site\_2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopen accumulation on the blossom end of  
the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

ORIGIN  
Query Match 1.2%; Score 49; DB 2; Length 570;  
Best Local Similarity 51.6%; Pred. No. 0.034;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2850 TACACTCTATCGACTGTATGCTCGTGCCTCCGACCATTTCCCTAGTGA 2909  
DB 289 TAATTTATTTCTGAAGGAAATGTGTGTCGAAAAATCCATGCTTGATCCTGCTGA 348  
QY 2910 TATTCACGCGGTTGACAGCTCTTCAAGCAGAGCTCCAGATCTCAAGATGTAATCAT 2969  
DB 349 TATTCGTTTAAAGCTGTAATATGTTGAGCGCTCACCACATGTAATGTGTGT 408  
QY 2970 CTCTCTACTAAAGAGATGATACCGCTTGCTAAGACCTATCTGTTGAGACTACGACGG 3029  
DB 409 ATTCCCTCAGAAAAGAAAGACCTCATCCGAATGAATGTTCTGGAGTGATTTGATGG 468  
QY 3030 CGAATAGCGCTGGTCTGCTGGATCCGAGATGTC 3066  
DB 469 GGATATCTACTTGTGTTGCTGGATCAAGACATGATC 505

RESULT 7  
CK851141

LOCUS CK851141 684 bp mRNA linear EST 08-MAR-2004  
DEFINITION 11217 Stolon Solanum tuberosum cDNA, mRNA sequence.  
ACCESSION CK851141  
VERSION CK851141.1 GI:45239751  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 684)  
Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De  
Koeijer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.  
Generation of ESTs from stolon tissues of potato  
TITLE Unpublished (2004)  
JOURNAL Contact: Barry Flinn  
COMMENT The Canadian Potato Genome Project - BioPlantech  
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA  
Email: [Bflinn@bioplantech.nb.ca](mailto:Bflinn@bioplantech.nb.ca)  
Seq primer: T3.  
Location/Qualifiers  
1..684  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Shepody"  
/db\_xref="taxon:4113"  
/tissue\_type="Stolon"  
/lab\_host="X110-Gold"  
/clone\_lib="Stolon"  
/note="Vector: Bluescript II SK(+) XR, Site\_1: EcoRI;  
Site\_2: XhoI; supplier: Developmental series. Plants from  
pathogen-free Solanum tuberosum var. Shepody, clone 1756,  
nuclear stock were grown in a greenhouse under natural  
conditions. RNA was isolated from stolon tissue."

ORIGIN  
Query Match 1.2%; Score 49; DB 7; Length 684;  
Best Local Similarity 53.4%; Pred. No. 0.036;  
Matches 103; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2870 ATGTCTCGTGGCGGATCCCGACCCATTTCCCTAGATATCCACGGGTTGAGCAG 2929  
DB 171 ATGTGTTGTTGCAAAAATCCATGCTTCATCTGTAATTTCTGTTTAAAGGCTG 230  
QY 2930 TCTTCAAGCAGAGCTCCACAGTCTCCAGATGTAATCATCTTCTACTAAAGAGATG 2989  
DB 231 TAGATGTTCCAGGCTGACACACATGTAATGTTGTAATTTCCCGAAGAAAGAAAA 290  
QY 2990 TACCGTTGCTAAGAGCTATCTGTGTGAGACTACGACCGCATATGCGCTGCT 3049  
DB 291 GACCTCATCCGAATGAATGTTCTGGAGTGATTTGATGAGATATCTATTGTTGCT 350  
QY 3050 GGGATCCGAGAT 3062  
DB 351 GGGATCAAGACAT 363

RESULT 8  
CK854274  
LOCUS CK854274 737 bp mRNA linear EST 08-MAR-2004  
DEFINITION 15379 Stolon Solanum tuberosum cDNA, mRNA sequence.  
ACCESSION CK854274  
VERSION CK854274.1 GI:45242874  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 737)  
Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De  
Koeijer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.  
Generation of ESTs from stolon tissues of potato  
TITLE Unpublished (2004)



Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

Location/Qualifiers  
1..925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPCI-98"  
/note="Tend : TET3"

## ORIGIN

Query Match 1.1%; Score 47; DB 9; Length 925;  
Best Local Similarity 16.2%; Pred. No. 0.15;  
Matches 51; Conservative 142; Mismatches 119; Indels 3; Gaps 1;  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

RESULT 11  
AM649811 597 bp mRNA linear EST 18-MAY-2001  
LOCUS EST328265 tomato germinating seedlings, TAMU Lycopersicon  
DEFINITION esculentum cDNA clone cLEI9L3 5', mRNA sequence.  
ACCESSION AM649811  
VERSION AM649811.1 GI:7411049  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 597)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
Niernan,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and  
Tankley,S.D.  
Generation of ESTs from germinating tomato seed  
Unpublished (2000)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.  
Location/Qualifiers  
1..597  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEI9L3"  
/issue\_type="whole seedlings"  
/dev\_stage="7 days post imbibition"  
/clone\_lib="tomato germinating seedlings, TAMU"  
/note="Vector: pBluescript SK(-); Site: 1; EcoRI; Site 2:  
XhoI; 7 days post imbibition on water-agar. Mixed stage  
whole germinating seedlings from seed coat emergence up  
to two centimeters in seeds not showing obvious signs of  
germination were discarded."

## ORIGIN

Query Match 1.1%; Score 45.6; DB 2; Length 597;  
Best Local Similarity 50.9%; Pred. No. 0.34;  
Matches 108; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

RESULT 12  
BF112386 367 bp mRNA linear EST 18-MAY-2001  
LOCUS EST439976 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION CLEGA1011 5' sequence, mRNA sequence.  
ACCESSION BF112386  
VERSION BF112386.1 GI:10942076  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 367)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tankley,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
Unpublished (2000)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>

```

/lab host="SOLR"
/clone.lib="romato breaker fruit"
/notes="vector: plusascripctskmudapt: Site 1: Ecori;
Site 2: XhoI, supplier: Boyce Thompson Institute. Fruit
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."
```

**ORIGIN**

Query Match	1.1%	Score 45;	DB 2;	Length 367;
Best Local Similarity	54.5%;	Pred. No. 0.44;		
Matches 90;	Conservative 0;	Mismatches 75;	Indels 0;	Gaps 0;

Qy 2902 CCTGTGATATCCAAACGGGTCGACAGCTTCAAGCCACAGTCTCAAGAT 2961

Db 14 CCTGTGATATTCGTGTTTAAAGCGCTGAATGTTGACGCGCTGCACCATGTAAGT 73

Qy 2962 GTATCATCTCTCTACTACTAAGGAGATGACCGGTGGCTAAAGAGCTATCTGGGAGAC 3021

Db 74 TGTGTTGTAATCCCTCAGAAAGAAAGACCTATCCGATGATATGTTCTGGGAGTAT 133

0y 3022 TACACGCGGATATGACCCTGGGTCGTGCTGGATCCGAGATCGTC 3066

Db 134 TTGGATGGGGATATCTACTTTTGTTCCTGGGATCAAGACATGATC 178

RESULT 13	
A1896465	
LOCUS	
A1896465	
421 bp	
EDNA	
14:pear	
EST 18-MAY-2001	

DEFINITION	CDNA clone
BST155836	tomato callus, TAMU lycopersicon esculentum
CLEC15G23, mRNA sequence.	
A1896465	
A1896465.1	GI:5602355
A1896465	

KEYWORDS  
SOURCE  
ORGANISM  
EST.  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

REFERENCE  
Alcala, J., Vreblov, J., White, R., Matern, A. L., Vision, T.,  
1 (pages 1 to 421)  
spermatophyta; Magnoliophyta; eudicotyledons; core eudots;  
asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
Alcala, J., Vreblov, J., White, R., Matern, A. L., Vision, T.,  
1 (pages 1 to 421)

TITLE	Author
Generation of ESTs from tomato callus tissue	Rochester, L.W., Wang, F., O'Connell, V., Craker, M.B., Edmonath, C.L., Fu, S., Romling, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

COMMENT  
Info: 800.428.6363  
Contact: CUGI  
Clemson University  
Clemson University

**FEATURES**

Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence:  
Location/Qualifiers

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="FA96"

```

```
/clone="CLECI5G23"  
/rissue_type="cal1us"  
/dev_age="25-40 days old"
```

```

/clone_1ib="tomato callus, TAMU"
note="vector: pBluescript
XhoI: supplier: Giovanni laboratory: clrc - Cotyledons
of seedlings 7 to 10 days post germination were avoided

```

undifferentiated masses. Tomato Callus Est Library"

**ORIGIN**

Query Match	1.1%;	Score 45;	DB 1;	Length 421;
Best Local Similarity	54.5%;	Pred. No. 0.46;		
Matches	90;	Conservative	0;	Mismatches 75; Indels 0; Gaps 0;

QY 2902 CCTAGTATATCCAAAGGTTGAGACGTCCTCAAGCAGAGCTCCACAGTCCAAAGAT 2961  
Db 14 CCTGGTATATTCGTGTTTAAAGCGCTGAATGTTGAGCGCTGCACACCATGTAAT 73

2962 GATATCATCTTCTCTACTAAAGAGATGTACCGCTTGTAAAGATCTTCTGTGAAC 3021  
 3022  
 74 TGGCTGTATTCCCTCAGAAAGAAAAGACCTCATCGAATGAATCTTCTGGAGATGAT 133

QY 3022 TACGACGGCGATATGAGCTGGGTCTGCTGGGATCCGAGATCCTTC 3066

Db 134 TTGCATGGGGATATCTCACTTTTGTTCGCGGATCAAGCATGATTC 178

RESULT 14  
AW933692  
LOCUS  
AW933692  
525 bp  
mpna  
linear  
EST 18-MAY-2001

DEFINITION	ESD5159533 tomato fruit mature green, TMU Lycopersicon esculentum
ACCESSION	CSNA clone cLEP56C18 5', mRNA sequence
VERSION	AW933692
ACCESSION	GI:8109093
VERSION	AW933692.1

KEYWORDS  
SOURCE  
ORGANISM  
EST.  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

REFERENCE  
AUTHORS  
Alcala, J., Vreblov, J., White, R., Matern, A. L., Holt, I. E., Liang, F.,  
1 (bases 1 to 525)  
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 525)

TITLE  
Generation of ESTs from tomato fruit tissue

COMMENT: Contact: CUGI  
Clemson University  
Clemson University

FEATURES

Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA96"

```

```

/clone="clEP56C18"
/ripen_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"

```

/clone\_1lb="tomato fruit mature green, TAMU"  
note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; cLEP - Fruit were tagged at the 1cm stage and  
harvested 3-5 days prior to ripening. Fruit were cut in

ORIGIN  
half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

Query Match	1.1%	Score 45;	DB 2;	Length 525;
Best Local Similarity	54.5%;	Pred. No. 0.49;		
Matches 90;	Conservative 0;	Mismatches 75;	Indels 0;	Gaps 0;

Oy	2902	CCTAGTGTATATCCAAACGGGTTTCGAGCAGTCTTTCAAAGCCAGAGCTCCACAGTCTCAAGGAT	2961
Db	14	CCGTGTGTATTCGTGTTTAAAGCTGTAAATGTTTCGAGCCGTGCACCCACATGTGTAT	73

QY 2962 GTAATCATCTTCTCTACTAAAGAGATGATCCGCTTGCTAAGACGTAATCTGTGAGAC 3021

Db 74 TGTGTTGTAATTCCTCGAAGAAAGAAAAGACCTCATCTCGAATGAATGTTCTGGAGAT 133

Dib

QY 3022 TACGACGGCATATGCGCTGGCTGCTGGATCCGAGATCGTC 3066  
 Db 134 TTGGATGGGATATCTACTTTGTTGCTGGATCAAGACATGATC 178

## RESULT 15

CF820058 742 bp mRNA linear EST 01-APR-2004  
 LOCUS EST697440 Coccidioides posadasii saprobic phase cDNA library, 2 to  
 DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDA224 3' end, mRNA  
 sequence.

ACCESSION CF820058  
 VERSION CF820058.1 GI:45926115  
 KEYWORDS EST.  
 SOURCE Coccidioides posadasii  
 ORGANISM Coccidioides posadasii  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 Onygenales; Mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 742)  
 AUTHORS Gardner M.J. and Cole G.T.  
 TITLE Analysis of gene expression in Coccidioides posadasii mycelia and  
 spherules via expressed sequence tags

JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST697441  
 Contact: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@igf.org

## FEATURES

source  
 Location/Qualifiers  
 1..742  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone="CIDA224"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_1lb="Coccidioides posadasii saprobic phase cDNA  
 library, 2 to 4 kb"  
 /note="Vector: pExpress 1; Site\_1: Not I; Site\_2: Eco RV;  
 Coccidioides posadasii saprobic phase cDNA library, size  
 fractionated cDNA 2 to 4 kb"

## ORIGIN

Query Match 1.1%; Score 44.8; DB 7; Length 742;  
 Best Local Similarity 52.1%; Pred. No. 0.63;  
 Matches 100; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 2874 CCTGTGCGCGGATCCCGACCATTCCTAGTATATCCAAAGGTTGACAGTCTT 2933  
 Db 46 CATATTTGCCGAATCGTCTTGATCCGGCGACATCCGGTGTAAAGTGTCAA 105  
 QY 2934 CAAGCCAGAGCTTCACAGTCTCAAGATGTAATCATCTTCTACTAAGAGATGACC 2993  
 Db 106 TGCACCGCGCTTACACCATCTTAGAGATGTATGTATCCACAGACAGAGATAGGA 165  
 QY 2994 GCTTGTAAAGACTATCTGTGAGAGCTACGACGGGATATAGCCTGGTCTGTGGGA 3053  
 Db 166 TATCCCAAGCATGTTGGGTGTGACTGTGATGATGATTAATCCTGTCAATTGGGA 225  
 QY 3054 TCCGAGATCGT 3065  
 Db 226 CCAAGATCTCCT 237

Search completed: August 11, 2005, 04:59:29  
 Job time : 12585 secs